

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE  
UNIVERSITÉ DU QUÉBEC

THÈSE PAR ARTICLES PRÉSENTÉE À  
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE  
L'UNIVERSITÉ DE NANTES  
(COTUTELLE)

COMME EXIGENCE PARTIELLE  
À L'OBTENTION DU  
DOCTORAT EN GÉNIE  
Ph.D.

COTUTELLE FRANCE-QUÉBEC

PAR  
Fabien DUBOT

DÉVELOPPEMENT DE TECHNIQUES DE DIAGNOSTIC NON INTRUSIF PAR  
TOMOGRAPHIE OPTIQUE

MONTREAL, LE 7 AOÛT 2015



Fabien Dubot, 2015



Cette licence Creative Commons signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette oeuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'oeuvre n'ait pas été modifié.

## **PRÉSENTATION DU JURY**

CETTE THÈSE A ÉTÉ ÉVALUÉE

PAR UN JURY COMPOSÉ DE:

M. Daniel Rousse, directeur de thèse  
Département de génie mécanique, École de Technologie Supérieure

M. Benoit Rousseau, directeur de thèse  
Laboratoire de Thermocinétique de Nantes, Université de Nantes

M. Yann Favennec, codirecteur  
Laboratoire de Thermocinétique de Nantes, Université de Nantes

M. Ambrish Chandra, président du jury  
Département de génie électrique, École de Technologie Supérieure

M. Denis Lemonnier, examinateur externe  
Institut Pprime, École Nationale Supérieure de Mécanique et d'Aérotechnique

M. Abdellatif El Badia, examinateur externe  
Laboratoire de Mathématiques Appliquées, Université de Technologie de Compiègne

M. Stéphane Hallé, membre du jury  
Département de génie mécanique, École de Technologie Supérieure

ELLE A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 12 JUIN 2015

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE





## REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de thèse québécois, Monsieur Daniel Rousse, pour la confiance qu'il m'a accordée en acceptant de financer et d'encadrer ce travail doctoral, ainsi que pour sa grande disponibilité malgré son emploi du temps très chargé. Ses qualités humaines et scientifiques exceptionnelles, son ouverture d'esprit et son sens aigu de la métaphore resteront toujours pour moi une source d'inspiration.

Je remercie également mon directeur de thèse français, Monsieur Benoit Rousseau, pour son soutien constant et ses conseils avisés. Son optimisme a été pour moi une source d'énergie importante dans l'aboutissement de cette recherche.

Je voudrais remercier tout particulièrement mon codirecteur de thèse, Monsieur Yann Favennec, pour m'avoir aidé dans l'accomplissement de ce travail. Merci pour son suivi, son aide, ses multiples encouragements ainsi que pour m'avoir fait partager son savoir dans le domaine des méthodes inverses.

Je tiens à exprimer toute ma gratitude aux membres du jury, Monsieur Ambrish Chandra, Monsieur Denis Lemonnier, Monsieur Abdellatif El Badia et Monsieur Stéphane Hallé pour avoir accepté de relire cette thèse et d'évaluer attentivement mon travail.

Je remercie vivement le Consulat Général de France à Québec à travers le Programme Frontenac de soutien à la mobilité des cotutelles France-Québec ainsi que Monsieur Michel Trottier pour le soutien financier dont j'ai pu bénéficier grâce à eux. J'adresse également mes remerciements aux partenaires financiers de la Chaire de recherche t3e, à savoir Ville de Lévis, Valéro Énergie, Ecosystem, CRE-CA, SDE-Lévis et Roche.

Je remercie de plus tous les développeurs des communautés du logiciel libre et de l'*open source*, notamment tous les contributeurs à Ubuntu, FreeFem++,  $\text{\LaTeX}$ , ParaView, Gnuplot, Imagemagick, Gmsh, dont j'ai utilisé le fruit de leur travail tout au long de cette thèse et grâce auxquels mes conditions de travail ont été des plus agréables.

Je tiens également à mentionner le plaisir que j'ai eu à travailler autant au sein du Laboratoire de Thermocinétique de Nantes que dans les locaux de la Chaire t3e, et j'en remercie ici tous les membres. Un merci particulier à Geneviève pour sa sympathie et son efficacité dans l'organisation et la résolution des problèmes administratifs auxquels j'ai pu être confronté au cours de ces dernières années.

Enfin, mes plus profonds remerciements vont à mes parents, et plus généralement à toute ma famille, mamies, papys, pour m'avoir donné la chance de faire de longue étude et pour m'avoir toujours soutenu, encouragé et aidé. Merci également à mes sœurs pour leur bonne humeur et pour les bons moments passés ensemble.

# DÉVELOPPEMENT DE TECHNIQUES DE DIAGNOSTIC NON INTRUSIF PAR TOMOGRAPHIE OPTIQUE

Fabien DUBOT

## RÉSUMÉ

Que ce soit dans les domaines des procédés industriels ou de l'imagerie médicale, on a assisté ces deux dernières décennies à un développement croissant des techniques optiques de diagnostic. L'engouement pour ces méthodes repose principalement sur le fait qu'elles sont totalement non invasives, qu'elles utilisent des sources de rayonnement non nocives pour l'homme et l'environnement et qu'elles sont relativement peu coûteuses et faciles à mettre en œuvre comparées aux autres techniques d'imagerie. Une de ces techniques est la Tomographie Optique Diffuse (TOD). Cette méthode d'imagerie tridimensionnelle consiste à caractériser les propriétés radiatives d'un Milieu Semi-Transparent (MST) à partir de mesures optiques dans le proche infrarouge obtenues à l'aide d'un ensemble de sources et détecteurs situés sur la frontière du domaine sondé. Elle repose notamment sur un modèle direct de propagation de la lumière dans le MST, fournissant les prédictions, et un algorithme de minimisation d'une fonction de coût intégrant les prédictions et les mesures, permettant la reconstruction des paramètres d'intérêt. Dans ce travail, le modèle direct est l'approximation diffuse de l'équation de transfert radiatif dans le régime fréquentiel tandis que les paramètres d'intérêt sont les distributions spatiales des coefficients d'absorption et de diffusion réduit.

Cette thèse est consacrée au développement d'une méthode inverse robuste pour la résolution du problème de TOD dans le domaine fréquentiel. Pour répondre à cet objectif, ce travail est structuré en trois parties qui constituent les principaux axes de la thèse.

Premièrement, une comparaison des algorithmes de Gauss-Newton amorti et de Broyden-Fletcher-Goldfarb-Shanno (BFGS) est proposée dans le cas bidimensionnel. Deux méthodes de régularisation sont combinées pour chacun des deux algorithmes, à savoir la réduction de la dimension de l'espace de contrôle basée sur le maillage et la régularisation par pénalisation de Tikhonov pour l'algorithme de Gauss-Newton amorti, et les régularisations basées sur le maillage et l'utilisation des gradients de Sobolev, uniformes ou spatialement dépendants, lors de l'extraction du gradient de la fonction coût, pour la méthode BFGS. Les résultats numériques indiquent que l'algorithme de BFGS surpasse celui de Gauss-Newton amorti en ce qui concerne la qualité des reconstructions obtenues, le temps de calcul ou encore la facilité de sélection du paramètre de régularisation.

Deuxièmement, une étude sur la quasi-indépendance du paramètre de pénalisation de Tikhonov optimal par rapport à la dimension de l'espace de contrôle dans les problèmes inverses d'estimation de fonctions spatialement dépendantes est menée. Cette étude fait suite à une observation réalisée lors de la première partie de ce travail où le paramètre de Tikhonov, déterminé par la méthode «L-curve», se trouve être indépendant de la dimension de l'espace de

contrôle dans le cas sous-déterminé. Cette hypothèse est démontrée théoriquement puis vérifiée numériquement sur un problème inverse linéaire de conduction de la chaleur puis sur le problème inverse non-linéaire de TOD. La vérification numérique repose sur la détermination d'un paramètre de Tikhonov optimal, défini comme étant celui qui minimise les écarts entre les cibles et les reconstructions. La démonstration théorique repose sur le principe de Morozov (*discrepancy principle*) dans le cas linéaire, tandis qu'elle repose essentiellement sur l'hypothèse que les fonctions radiatives à reconstruire sont des variables aléatoires suivant une loi normale dans le cas non-linéaire. En conclusion, la thèse démontre que le paramètre de Tikhonov peut être déterminé en utilisant une paramétrisation des variables de contrôle associée à un maillage lâche afin de réduire les temps de calcul.

Troisièmement, une méthode inverse multi-échelle basée sur les ondelettes associée à l'algorithme de BFGS est développée. Cette méthode, qui s'appuie sur une reformulation du problème inverse original en une suite de sous-problèmes inverses de la plus grande échelle à la plus petite, à l'aide de la transformée en ondelettes, permet de faire face à la propriété de convergence locale de l'optimiseur et à la présence de nombreux minima locaux dans la fonction coût. Les résultats numériques montrent que la méthode proposée est plus stable vis-à-vis de l'estimation initiale des propriétés radiatives et fournit des reconstructions finales plus précises que l'algorithme de BFGS ordinaire tout en nécessitant des temps de calcul semblables.

Les résultats de ces travaux sont présentés dans cette thèse sous forme de quatre articles. Le premier article a été accepté dans l'*International Journal of Thermal Sciences*, le deuxième est accepté dans la revue *Inverse Problems in Science and Engineering*, le troisième est accepté dans le *Journal of Computational and Applied Mathematics* et le quatrième a été soumis au *Journal of Quantitative Spectroscopy & Radiative Transfer*. Dix autres articles ont été publiés dans des comptes-rendus de conférences avec comité de lecture. Ces articles sont disponibles en format pdf sur le site de la Chaire de recherche t3e ([www.t3e.info](http://www.t3e.info)).

**Mots clés:** problème inverse, transfert radiatif, optimisation, régularisation, ondelette, gradients de Sobolev

# **DEVELOPMENT OF NON-INVASIVE DIAGNOSTIC TECHNIQUES BY OPTICAL TOMOGRAPHY**

Fabien DUBOT

## **ABSTRACT**

Whether in the fields of industrial process or medical imaging, there has been an increasing development of optical diagnostic techniques over the past two decades. Enthusiasm for these methods mainly relies on the fact that they are non-intrusive, they use radiation sources safe to human and environment, and they are relatively inexpensive and easy to implement compared to other imaging techniques. One of these techniques is the Diffuse Optical Tomography (DOT). This three-dimensional imaging method consists in characterizing radiative properties of a participating medium from optical measurements in near-infrared provided by a set of sources and sensors located on the frontier of the probed medium. It is especially based on a forward model of light propagation in the participating medium, providing the predictions, and a minimization algorithm of a cost function which integrates the predictions and measurements, allowing the reconstruction of parameters of interest. In this work, the forward model is the diffuse approximation of the radiative transfer equation in the frequency domain while parameters of interest are the spatial distributions of absorption and reduced scattering coefficients.

The objective of the work presented in this thesis is the development of a robust inverse method for the solution of the DOT problem in the frequency domain. To meet this objective, this work is structured in three main parts.

First, a comparison between the damped Gauss-Newton and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithms is proposed in the two-dimensional case. Two regularization methods are combined for each of the two algorithms, namely the mesh-based reduction of the control space dimension and the Tikhonov penalization for the damped Gauss-Newton algorithm, and the regularizations respectively based on the mesh and the use of Sobolev gradients, uniform or spatially dependent, when extracting the cost function gradient, for the BFGS method. Numerical results indicate that the BFGS algorithm outperforms the damped Gauss-Newton algorithm in term of the quality of the obtained reconstructions, CPU time and the ease of selecting the regularization parameter.

Second, a study on the quasi-independence of the optimal Tikhonov penalization parameter with respect to the dimension of the control space in inverse problems of space-dependent function estimation is performed. This study follows an observation made in the first part of this work where the Tikhonov parameter, determined by the “L-curve” method, is found to be independent of the dimension of the control space in the under-determined case. This assumption is demonstrated theoretically and verified numerically on a linear inverse heat conduction problem then on the non-linear inverse problem of the DOT. The numerical verification is based on the determination of an optimal Tikhonov parameter, defined as the one which minimizes

the misfit between the target and the reconstruction. The theoretical demonstration relies on the discrepancy principle in the linear case, while it mainly relies on the assumption that radiative functions to be reconstructed are normally distributed random variables in the non-linear case. In conclusion, this thesis demonstrates that the Tikhonov parameter can be determined using a control space parameterization associated with a coarse mesh in order to lower computational time.

Third, a wavelet multi-scale method associated with the BFGS algorithm is developed. This method, which relies on a reformulation of the original inverse problem into a sequence of sub-inverse problems of different scales using wavelet transform, from the largest scale to the smallest one, enables to cope with the local convergence property of the optimizer and the presence of numerous local minima in the cost function. Numerical results show that the proposed method is more stable with respect to the initial guess of the radiative properties and provides more accurate final reconstructions compared to the ordinary BFGS algorithm while requiring similar computational times.

The thesis is presented in the form of four articles. The first article is accepted in the *International Journal of Thermal Sciences*, the second is accepted in the journal *Inverse Problems in Science and Engineering*, the third is accepted in the *Journal of Computational and Applied Mathematics* and the fourth has been submitted in the *Journal of Quantitative Spectroscopy & Radiative Transfer*. Ten other articles have been published in conferences with reviewing committee. Pdf versions of these papers are available on the web site of the industrial research chair in technologies of energy and energy efficiency (t3e) ([www.t3e.info](http://www.t3e.info)).

**Keywords:** inverse problem, radiative transfer, optimization, regularization, wavelet, Sobolev gradients

## TABLE DES MATIÈRES

	Page
INTRODUCTION .....	1
CHAPITRE 1 REVUE DE LITTÉRATURE .....	9
1.1 Grandeurs photométriques .....	10
1.1.1 Luminance .....	10
1.1.2 Vecteur flux radiatif .....	11
1.1.3 Densité volumique d'énergie .....	11
1.1.4 Densité de photons .....	11
1.2 Grandeurs caractéristiques du transport de la lumière .....	12
1.2.1 Indice de réfraction .....	12
1.2.2 Absorption .....	12
1.2.3 Émission .....	13
1.2.4 Diffusion de la lumière .....	13
1.2.5 Fonction de phase .....	14
1.3 L'équation de transfert radiatif .....	16
1.3.1 Dérivation de l'équation de transfert radiatif selon l'approche phénoménologique .....	16
1.3.2 Conditions de bord .....	17
1.4 De l'équation de transfert radiatif à l'approximation diffuse .....	20
1.4.1 Équation sur la densité de photons .....	20
1.4.2 Équation sur le vecteur flux radiatif .....	21
1.4.3 Conditions de bord .....	24
1.4.4 Quantité mesurable .....	27
1.4.5 Modélisation des sources .....	28
1.4.6 Passage au régime fréquentiel .....	30
1.4.7 Domaine de validité .....	32
CHAPITRE 2 ARTICLE 1 : REGULARIZATION OPPORTUNITIES FOR THE DIFFUSE OPTICAL TOMOGRAPHY PROBLEM .....	35
2.1 Introduction .....	36
2.2 Problem statement .....	39
2.2.1 Forward model .....	39
2.2.2 Cost function .....	42
2.3 Mathematical settings .....	43
2.4 Optimization algorithms .....	46
2.4.1 Gauss-Newton algorithm .....	46
2.4.2 Broyden-Fletcher-Goldfarb-Shanno algorithm .....	50
2.5 Mixing regularization tools for Gauss-Newton and BFGS algorithms .....	54
2.5.1 Test presentation .....	54
2.5.2 Definition of reconstruction errors .....	55

2.5.3	Gauss-Newton algorithm : mixing Tikhonov penalization and mesh-based parameterization .....	57
2.5.4	BFGS algorithm : comparison between adimensionalization and gradient scaling strategies .....	60
2.5.5	BFGS algorithm : mesh-based parameterization .....	63
2.5.6	BFGS algorithm : mixing mesh-based parameterization and Sobolev gradients method .....	65
2.5.7	BFGS algorithm : space-dependent Sobolev gradients method .....	70
2.6	Conclusion .....	72
CHAPITRE 3 ARTICLE 2 : QUASI-OPTIMAL TIKHONOV PENALIZATION AND PARAMETERIZATION COARSENESS IN SPACE-DEPENDENT FUNCTION ESTIMATION .....		
3.1	Introduction .....	79
3.2	Space-dependent heat flux estimation .....	83
3.2.1	Problem statement .....	83
3.2.2	Numerical results .....	85
3.3	Optical tomography .....	87
3.3.1	Forward model and inversion setting .....	87
3.3.2	Cost function derivative .....	89
3.3.3	Parameterization .....	90
3.3.4	Tikhonov regularization .....	92
3.3.5	Numerical validation .....	94
3.4	Conclusion .....	100
CHAPITRE 4 ARTICLE 3 : A WAVELET MULTI-SCALE METHOD FOR THE INVERSE PROBLEM OF DIFFUSE OPTICAL TOMOGRAPHY .....		
4.1	Introduction .....	102
4.2	The diffuse optical tomography problem .....	105
4.3	Cost function and its gradients .....	106
4.4	Optimization .....	108
4.5	Wavelet multi-scale method .....	110
4.5.1	Orthogonal wavelet bases .....	110
4.5.2	Multi-scale reconstruction method .....	113
4.6	Numerical results .....	115
4.6.1	Tests presentation .....	115
4.6.2	Definition of reconstruction errors .....	117
4.6.3	Test 1 : reconstruction of discontinuous target properties .....	118
4.6.4	Test 2 : reconstruction of continuous target properties .....	121
4.6.5	Discussion .....	124
4.7	Conclusion .....	126



CHAPITRE 5	ARTICLE 4 : SPACE-DEPENDENT SOBOLEV GRADIENTS AS A REGULARIZATION FOR INVERSE RADIATIVE TRANSFER PROBLEMS .....	131
5.1	Introduction .....	131
5.2	Space-dependent Sobolev cost function gradient .....	135
5.3	Application on 2D RTE-based optical tomography .....	141
5.3.1	Inverse problem statement .....	141
5.3.2	Cost function gradient derivation .....	143
5.3.3	Numerical results .....	145
5.4	Application on 3D DA-based optical tomography .....	147
5.4.1	Inverse problem statement .....	147
5.4.2	Cost function gradients derivation .....	153
5.4.3	Numerical results .....	154
5.5	Conclusion .....	156
CHAPITRE 6	RÉSULTATS COMPLÉMENTAIRES .....	159
6.1	Méthode inverse multi-échelle basée sur les ondelettes de Haar pour le modèle direct des sources collimatées .....	159
6.2	<i>Crosstalk</i> entre les paramètres .....	163
6.3	Reconstructions tridimensionnelles .....	169
CHAPITRE 7	DISCUSSIONS DES RÉSULTATS .....	175
CONCLUSION ET PERSPECTIVES .....		179
ANNEXE I	CALCULATION OF THE COST FUNCTION GRADIENT WITH THE ADJOINT-STATE METHOD .....	185
ANNEXE II	PARAMÉTRISATION DES VARIABLES DE CONTRÔLE ET MÉTHODES DE RECHERCHE LINÉAIRE DANS UN CODE D'INVERSION DE L'APPROXIMATION DE DIFFUSION BASÉ SUR LE L-BFGS .....	191
ANNEXE III	NUMERICAL VALIDATION OF THE FORWARD MODEL SOLUTION .....	203
ANNEXE IV	LISTE DES PUBLICATIONS EN LIEN AVEC LA THÈSE .....	209
RÉFÉRENCES BIBLIOGRAPHIQUES .....		212



## LISTE DES TABLEAUX

		Page
TAB. 0.1	Comparaison de différentes méthodes de diagnostic. Adapté de (Charette <i>et al.</i> , 2011). ....	4
TAB. 1.1	Quelques valeurs du paramètre $A = \frac{1+R_{\text{eff}}}{1-R_{\text{eff}}}$ à l'interface entre deux milieux d'indices de réfraction différents selon (Groenhuis <i>et al.</i> , 1983), (Haskell <i>et al.</i> , 1994) et (Keijzer <i>et al.</i> , 1988) ( $n_a = 1$ ). ....	27
TAB. 2.1	Gauss-Newton algorithm. Obtained reconstruction errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for 4 meshes and 3 levels of noise. ....	60
TAB. 2.2	BFGS algorithm; mesh $\mathcal{M}_1$ . Obtained reconstruction errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for both the gradient scaling approach and the adimensionalisation approach. 3 levels of noise are considered. ....	63
TAB. 2.3	BFGS algorithm. Obtained reconstruction errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for 4 meshes and 3 levels of noise. ....	66
TAB. 2.4	BFGS algorithm; mesh $\mathcal{M}_3$ . Obtained reconstruction errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for 4 values for the Sobolev regularization parameter. 3 levels of noise are considered. ....	70
TAB. 3.1	Value of the Tikhonov parameter $\vartheta_{\text{dp}}^*$ solution of (3.6) for different discretizations and different noise magnitudes. ....	86
TAB. 3.2	Dimensions of finite element spaces for $\varphi$ , $\check{\varphi}$ and $\lambda$ . ....	95
TAB. 3.3	CPU time comparisons for the generation of the error curves $E(\vartheta)$ . ....	97
TAB. 4.1	Errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for the test 1 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the ordinary L-BFGS method. ....	119
TAB. 4.2	Errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for the test 1 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the Haar wavelet. ....	120
TAB. 4.3	Errors $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ for the test 1 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D4 wavelet. ....	120

TAB. 4.4	Errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for the test 1 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D6 wavelet. ....	120
TAB. 4.5	Errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for the test 2 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the ordinary L-BFGS method. A cross indicates the divergence of the algorithm.....	124
TAB. 4.6	Errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for the test 2 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the Haar wavelet.....	124
TAB. 4.7	Errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for the test 2 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D4 wavelet. A cross indicates the divergence of the algorithm. ....	125
TAB. 4.8	Errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for the test 2 and different values of the a priori functions $\kappa_{ap}$ and $\sigma_{ap}$ ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D6 wavelet. A cross indicates the divergence of the algorithm. ....	125

## TABLE DES FIGURES

		Page
FIG. 0.1	Appareillage de TOD pour la détection de tumeurs dans le sein. Tiré de (Hajihashemi <i>et al.</i> , 2012). . . . .	4
FIG. 0.2	Représentation de la fenêtre thérapeutique : coefficient d'absorption des principaux composants absorbeurs dans les tissus biologiques en fonction de la longueur d'onde. Extrait de (Hillman, 2002). . . . .	6
FIG. 0.3	Schématisation du principe général du problème de tomographie optique. . . . .	7
FIG. 1.1	Géométrie utilisée pour la définition de la luminance. . . . .	10
FIG. 1.2	Relation entre les rayonnements incident $s'$ , réfléchi $s$ et réfracté $s''$ pour un rayonnement incident quittant le milieu. . . . .	19
FIG. 1.3	Relation entre les rayonnements incident $s^*$ , réfléchi $s'$ et réfracté $s$ pour un rayonnement incident entrant dans le milieu. . . . .	19
FIG. 2.1	Synthetic data mesh : (a) $\mathcal{M}_s$ , 13 143 nodes ; (b) $\mathcal{M} = \mathcal{M}_1$ , 3 764 nodes ; (c) $\mathcal{M}_2$ , 1 878 nodes ; (d) $\mathcal{M}_3$ , 978 nodes ; (e) $\mathcal{M}_4$ , 170 nodes ; (f) locations of sources and sensors ; (g) maps of $\kappa_t$ ; (h) maps of $\sigma_t$ . . . . .	56
FIG. 2.2	L-curves (left) and $(\pi/4)$ -rotated corresponding curves (right) for a 30 dB noise (top) and 10 dB noise (bottom). All curves are built for the three meshes $\mathcal{M}_1$ , $\mathcal{M}_2$ and $\mathcal{M}_3$ where $n_p = 3764, 1878$ and $978$ , respectively. Presented Tikhonov parameters are equal to 0.4, 0.6, 0.8, 0.9, 1, 1.1, 1.5, 2, 5, 10 and 100. . . . .	59
FIG. 2.3	Gauss-Newton algorithm ; 30 dB noise. Reconstructions for 4 control-space parameterizations. Left : $\kappa$ , right : $\sigma$ . First row : targets, next rows : reconstructions. . . . .	61
FIG. 2.4	Gauss-Newton algorithm ; 10 dB noise. Reconstructions for 4 control-space parameterizations. Left : $\kappa$ , right : $\sigma$ . First row : targets, next rows : reconstructions. . . . .	62
FIG. 2.5	BFGS algorithm ; mesh $\mathcal{M}_1$ ; 30 dB noise. Reconstructions for both the gradient scaling approach and the adimensionalisation approach. Left : $\kappa$ , right : $\sigma$ . First row : targets, second row : gradient scaling, third row : adimensionalization. . . . .	64

FIG. 2.6	BFGS algorithm ; mesh $\mathcal{M}_1$ ; 10 dB noise. Reconstructions for both the gradient scaling approach and the adimensionalisation approach. Left : $\kappa$ , right : $\sigma$ . First row : targets, second row : gradient scaling, third row : adimensionalization. ....	65
FIG. 2.7	BFGS algorithm ; 30 dB noise. Reconstructions for 4 control-space parameterizations. Left : $\kappa$ , right : $\sigma$ . First row : targets, next rows : reconstructions. ....	67
FIG. 2.8	BFGS algorithm ; 10 dB noise. Reconstructions for 4 control-space parameterizations. Left : $\kappa$ , right : $\sigma$ . First row : targets, next rows : reconstructions. ....	68
FIG. 2.9	BFGS algorithm ; mesh $\mathcal{M}_3$ , 30 dB noise. Obtained reconstruction errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for 4 values for the Sobolev regularization parameter. Left : $\kappa$ , right : $\sigma$ . First row : targets, next rows : reconstructions. ....	75
FIG. 2.10	BFGS algorithm ; mesh $\mathcal{M}_3$ , 10 dB noise. Obtained reconstruction errors $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ for 4 values for the Sobolev regularization parameter. Left : $\kappa$ , right : $\sigma$ . First row : targets, next rows : reconstructions. ....	76
FIG. 2.11	The space-dependent Sobolev filter $\vartheta_{25}^S(\mathbf{r})$ . ....	77
FIG. 2.12	Red and green points : evolution of errors $e_1^\kappa$ (a), $e_1^\sigma$ (b), $e_2^\kappa$ (c) and $e_2^\sigma$ (d) with respect to the regularization parameter $\nu$ . The parameter $\nu$ ranges from 1 to 100. Black line : errors $e_1^\kappa$ , $e_1^\sigma$ , $e_2^\kappa$ and $e_2^\sigma$ found in Table 2.3 (i.e. obtained with the BFGS method), and for 10 dB and the mesh $\mathcal{M}_3$ (i.e. corresponding to the best control-space parameterization). ....	77
FIG. 2.13	BFGS algorithm ; mesh $\mathcal{M}_3$ , 10 dB noise. Reconstructions with the space-dependent Sobolev regularization with $\nu = 25$ . Left : $\kappa$ , right : $\sigma$ . First row : targets, second row : reconstructions with the ordinary $L^2$ -inner product, third row : reconstructions with the $\vartheta_{25}^S(\mathbf{r})$ Sobolev filter. ....	78
FIG. 3.1	Distance from the solution to the target $E(\vartheta)$ as a function of the Tikhonov parameter $\vartheta$ , for three control space dimensions $\Xi$ equal to 7, 15 and 20. It is seen that $\vartheta^* = \arg \min E(\vartheta)$ is independent of $\Xi$ . The quasi-optimal Tikhonov parameter found according to the discrepancy principle has also been added. ....	87

FIG. 3.2	Reconstructions $\varphi_{\vartheta}(r_2)$ for $\Xi = 15$ (top), $\Xi = 10$ (middle) and $\Xi = 7$ (bottom) for under-regularization ( $\vartheta = 10^{-15}$ ), over-regularization ( $\vartheta = 1$ ) and appropriate regularization ( $\vartheta = 5 \cdot 10^{-10} \approx \vartheta^*$ ). Note that points for $\Xi = 15$ and $\Xi = 10$ (over-parameterization) with $\vartheta = 10^{-15}$ (under-regularization) are not presented because of divergence. The noise variance $\varepsilon_l^2 = 0.001$ was used after generating the synthetic data. ....	88
FIG. 3.3	Test medium geometry representation. Eight sources are located on the boundary. For each source (which constitutes a test), the emerging radiation is measured on all sensors.....	95
FIG. 3.4	Distance from the solution to the target $E(\vartheta)$ as a function of the Tikhonov parameter $\vartheta$ , for $(\kappa_{ap}, \sigma_{ap}) = (0.08, 20) \text{ cm}^{-1}$ and three control space dimensions $\Xi_{\lambda}$ equal to 384, 1382 and 2924. It is seen that $\vartheta^* = \arg \min E(\vartheta)$ is quasi-independent of $\Xi$ .....	97
FIG. 3.5	Reconstruction of $\kappa$ (left) and $\sigma$ (right) with $\vartheta = 0.2$ , $(\kappa_{ap}, \sigma_{ap}) = (0.08, 20) \text{ cm}^{-1}$ and $\Xi_{\lambda} = 2924$ . It should be noted the divergence for the reduced diffusion coefficient. ....	98
FIG. 3.6	Reconstruction of $\kappa$ (left) and $\sigma$ (right) with $\vartheta = 0.4$ , $(\kappa_{ap}, \sigma_{ap}) = (0.08, 20) \text{ cm}^{-1}$ and $\Xi_{\lambda} = 2924$ .....	98
FIG. 3.7	Reconstruction of $\kappa$ (left) and $\sigma$ (right) with $\vartheta = 1.4$ , $(\kappa_{ap}, \sigma_{ap}) = (0.088, 24) \text{ cm}^{-1}$ and $\Xi_{\lambda} = 2924$ . ....	99
FIG. 4.1	Test 1. Reconstruction of radiative properties with $\kappa_{ap} = 0.08$ and $\sigma_{ap} = 20$ . Left-hand side, results for $\kappa$ ; right-hand side, results for $\sigma$ . (a-b) target properties; (c-d) ordinary L-BFGS method; (e-f) Haar wavelet.....	121
FIG. 4.2	Test 1. Reconstruction of radiative properties with $\kappa_{ap} = 0.08$ and $\sigma_{ap} = 20$ . Left-hand side, results for $\kappa$ ; right-hand side, results for $\sigma$ . (a-b) target properties; (c-d) D4 wavelet; (e-f) D6 wavelet. ....	122
FIG. 4.3	Test 2. Reconstruction of radiative properties with $\kappa_{ap} = 0.1$ and $\sigma_{ap} = 22.5$ . Left-hand side, results for $\kappa$ ; right-hand side, results for $\sigma$ . (a-b) target properties; (c-d) ordinary L-BFGS method; (e-f) Haar wavelet.....	126
FIG. 4.4	Test 2. Reconstruction of radiative properties with $\kappa_{ap} = 0.1$ and $\sigma_{ap} = 22.5$ . Left-hand side, results for $\kappa$ ; right-hand side, results for $\sigma$ . (a-b) target properties; (c-d) D4 wavelet; (e-f) D6 wavelet. ....	127

FIG. 4.5	Test 2. Reconstruction of radiative properties with the Haar wavelet. Left-hand side, results for $\kappa$ ; right-hand side, results for $\sigma$ . (a-b) target properties; (c-d) $\kappa_{ap} = 0.08$ and $\sigma_{ap} = 20$ ; (e-f) $\kappa_{ap} = 0.08$ and $\sigma_{ap} = 22.5$ . ....	128
FIG. 5.1	Schematic description of the experiment. Only one source is on at once, while others are off. For each source configuration number $k = 1, \dots, K$ , the emerging intensity is measured on all sensors $d = 1, \dots, D$ in different directions $r = 1, \dots, R$ . ....	142
FIG. 5.2	Evolution of the cost function with respect to iterations, for $v = +\infty$ , 200, 100 and 50 and a 20 dB, 15 dB and 10 dB noise. The algorithms stops at iteration $k$ when $\frac{ j^k - j^{k-1} }{j^{k-1}} < 10^{-2}$ . ....	147
FIG. 5.3	20 dB. Evolution of errors $e_1$ (top) and $e_2$ (bottom) for $\kappa$ (left) and $\sigma_s$ (right), with respect to iterations, for $v = +\infty$ , 200, 100 and 50. The algorithms stops at iteration $k$ when $\frac{ j^k - j^{k-1} }{j^{k-1}} < 10^{-2}$ . ....	148
FIG. 5.4	15 dB. Evolution of errors $e_1$ (top) and $e_2$ (bottom) for $\kappa$ (left) and $\sigma_s$ (right), with respect to iterations, for $v = +\infty$ , 200, 100 and 50. The algorithms stops at iteration $k$ when $\frac{ j^k - j^{k-1} }{j^{k-1}} < 10^{-2}$ . ....	149
FIG. 5.5	10 dB. Evolution of errors $e_1$ (top) and $e_2$ (bottom) for $\kappa$ (left) and $\sigma_s$ (right), with respect to iterations, for $v = +\infty$ , 200, 100 and 50. The algorithms stops at iteration $k$ when $\frac{ j^k - j^{k-1} }{j^{k-1}} < 10^{-2}$ . ....	150
FIG. 5.6	20 dB. Reconstructions obtained for $\kappa$ (left) and $\sigma_s$ (right). 1st row : targets, then : $v = +\infty$ , 200, 100 and 50, respectively. ....	151
FIG. 5.7	10 dB. Reconstructions obtained for $\kappa$ (left) and $\sigma_s$ (right). 1st row : targets, then : $v = +\infty$ , 200, 100 and 50, respectively. ....	152
FIG. 5.8	Synthetic data mesh : 132651 nodes (left); mesh of the state and adjoint variable : 68921 nodes (center); mesh associated to radiative properties $\kappa$ and $\sigma$ : 9261 nodes (right). ....	155
FIG. 5.9	10 dB. Evolution of errors $e_1$ (top) and $e_2$ (bottom) for $\kappa$ (left) and $\sigma$ (right), with respect to iterations, for $v = +\infty$ , 70, 40 and 10. The algorithms stops at iteration $k$ when $\frac{ j^k - j^{k-1} }{j^{k-1}} < 10^{-3}$ . ....	156
FIG. 5.10	10 dB. Reconstructions obtained in the plane passing through the center of the cube of which the normal vector $\mathbf{n}$ equals $(1, -1, 0)^\top$ ,	



	for $\kappa$ (left) and $\sigma$ (right). 1st row : targets, then : $v = +\infty, 70, 40$ and 10, respectively. ....	157
FIG. 6.1	Configuration des tests des sections 6.1 et 6.2. ....	160
FIG. 6.2	Reconstructions $\kappa$ et $\sigma$ . SNR = 20 dB. $\kappa_{ap} = 0,08$ , $\sigma_{ap} = 20$ . 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS, maillage composé de $128 \times 128$ nœuds. 3 <sup>e</sup> ligne : méthode de BFGS, maillage composé de $80 \times 80$ nœuds. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	162
FIG. 6.3	Reconstructions $\kappa$ et $\sigma$ . SNR = 10 dB. $\kappa_{ap} = 0,08$ , $\sigma_{ap} = 20$ . 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS, maillage composé de $128 \times 128$ nœuds. 3 <sup>e</sup> ligne : méthode de BFGS, maillage composé de $80 \times 80$ nœuds. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	163
FIG. 6.4	Reconstructions $\kappa$ et $\sigma$ . SNR = 20 dB. $\kappa_{ap} = 0,08$ , $\sigma_{ap} = 20$ . Maillage composé de $80 \times 80$ nœuds. 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS. 3 <sup>e</sup> ligne : méthode multi-échelle. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	164
FIG. 6.5	Reconstructions $\kappa$ et $\sigma$ . SNR = 10 dB. $\kappa_{ap} = 0,08$ , $\sigma_{ap} = 20$ . Maillage composé de $80 \times 80$ nœuds. 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS. 3 <sup>e</sup> ligne : méthode multi-échelle. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	165
FIG. 6.6	Reconstructions $\kappa$ et $\sigma$ . SNR = 20 dB. $\kappa_{ap} = 0,084$ , $\sigma_{ap} = 21$ . Maillage composé de $80 \times 80$ nœuds. 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS. 3 <sup>e</sup> ligne : méthode multi-échelle. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	166
FIG. 6.7	Reconstructions $\kappa$ et $\sigma$ . SNR = 10 dB. $\kappa_{ap} = 0,084$ , $\sigma_{ap} = 21$ . Maillage composé de $80 \times 80$ nœuds. 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS. 3 <sup>e</sup> ligne : méthode multi-échelle. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	167
FIG. 6.8	Illustration du phénomène de <i>crosstalk</i> entre les paramètres (1). SNR = 20 dB. $\kappa_{ap} = 0,08$ , $\sigma_{ap} = 20$ . Maillage composé de $80 \times 80$ nœuds. 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS. 3 <sup>e</sup> ligne : méthode multi-échelle. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	170
FIG. 6.9	Illustration du phénomène de <i>crosstalk</i> entre les paramètres (2). SNR = 20 dB. $\kappa_{ap} = 0,08$ , $\sigma_{ap} = 20$ . Maillage composé de $80 \times 80$ nœuds. 1 <sup>re</sup> ligne : propriétés cibles. 2 <sup>e</sup> ligne : méthode de BFGS. 3 <sup>e</sup> ligne : méthode multi-échelle. Gauche : $\kappa$ ; Droite : $\sigma$ . ....	171

FIG. 6.10	(a) Géométrie utilisée pour le calcul des données synthétiques, contenant les deux inclusions ; (b) maillage de la variable d'état ; (c) maillage des variables de contrôle. ....	173
FIG. 6.11	Coupes diagonales des reconstructions des propriétés radiatives tridimensionnelles pour deux valeurs du paramètre de régularisation de Sobolev $\vartheta^S$ . Haut : coefficient d'absorption ; Bas : coefficient de diffusion réduit. ....	173
FIG. 7.1	Schématisation du principe de la méthode de fonctionnalisation des propriétés radiatives des matériaux. ....	183

## **LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES**

TO	Tomographie optique
OT	Optical tomography
TOD	Tomographie optique diffuse
DOT	Diffuse optical tomography
ÉTR	Équation de transfert radiatif
RTE	Radiative transfer equation
AD	Approximation diffuse
DA	Diffuse approximation
MST	Milieu semi-transparent
GN	Gauss-Newton
(L-)BFGS	(Limited-memory-)Broyden-Fletcher-Goldfarb-Shanno
CPU	Central processing unit
MUMPS	MUltifrontal Massively Parallel sparse direct Solver
SNR	Signal-to-noise ratio
TI-X	Tomographie informatisée par rayons-X
IRM	Imagerie par résonance magnétique
TEP	Tomographie par émission de positrons
IR	Infrarouge
NIR	Near-infrared
DEL	Diode électroluminescente



## LISTE DES SYMBOLES ET UNITÉS DE MESURE

$\mathbb{R}$	Ensemble des nombres réels
$\mathbb{C}$	Ensemble des nombres complexes
$\Omega$	Domaine d'étude, ouvert borné de $\mathbb{R}^{n_\Omega}$
$n_\Omega$	Dimension du domaine d'étude $\Omega$
$\partial\Omega$	Frontière du domaine d'étude $\Omega$
$\partial\Omega_s$	Emplacement des sources
$\partial\Omega_d$	Emplacement des détecteurs
$\mathcal{S}^{n_\Omega-1}$	Sphère unité de $\mathbb{R}^{n_\Omega}$
$\mathbf{r}$	Variable d'espace, $\mathbf{r} = (r_1, \dots, r_{n_\Omega})$
$s$	Variable angulaire
$t$	Variable de temps
$\mathbf{n}$	Vecteur normal unitaire
$\nu$	Fréquence du rayonnement électromagnétique, [s <sup>-1</sup> ]
$\omega$	Fréquence de modulation de la source lumineuse, [s <sup>-1</sup> ]
$c_0$	Célérité de la lumière dans le vide, [m s <sup>-1</sup> ]
$c$	Vitesse de la lumière dans $\Omega$ , [m s <sup>-1</sup> ]
$n_a$	Indice de réfraction du milieu ambiant
$n$	Indice de réfraction de $\Omega$
$h$	Constante de Planck, [J s]
$i$	Variable imaginaire
$L_\nu$	Luminance monochromatique à la fréquence $\nu$ , [W m <sup>-2</sup> sr <sup>-1</sup> Hz <sup>-1</sup> ]
$\mathbf{J}$	Vecteur flux radiatif, [W m <sup>-2</sup> ]
$U$	Densité volumique d'énergie, [J m <sup>-3</sup> ]
$\varphi$	Densité de photons, [W m <sup>-2</sup> ]

$\check{\phi}$	Données synthétiques, $[\text{W m}^{-2}]$
$\varphi^*$	Variable adjointe
$\kappa$	Coefficient d'absorption, $[\text{m}^{-1}]$
$\varkappa$	Coefficient d'absorption adimensionnel
$\sigma_s$	Coefficient de diffusion, $[\text{m}^{-1}]$
$\sigma$	Coefficient de diffusion réduit, $[\text{m}^{-1}]$
$\varsigma$	Coefficient de diffusion réduit adimensionnel
$\beta$	Coefficient d'extinction, $[\text{m}^{-1}]$
$\lambda$	Égal à $\kappa$ ou $\sigma$ , $[\text{m}^{-1}]$
$D$	Coefficient de diffusion macroscopique, $[\text{m}]$
$\Psi$	Fonction de phase
$g$	Facteur d'asymétrie
$\Phi_-$	Densité de flux radiatif entrante, $[\text{W m}^{-2}]$
$\Phi_+$	Densité de flux radiatif sortante, $[\text{W m}^{-2}]$
$R_1, R_2$	Facteurs de réflexion de Fresnel
$R_{\text{eff}}$	Coefficient de réflexion effectif non directionnel
$A$	Paramètre caractérisant la réflexion à la frontière du domaine
$f$	Densité de flux d'une source extérieure diffuse, $[\text{W m}^{-2}]$
$q$	Terme source volumique associé à une source extérieure collimatée, $[\text{W m}^{-3}]$
$\gamma$	Variable de contrôle du problème inverse
$\mathbb{1}$	Fonction indicatrice
$\delta(.)$	Distribution de Dirac
$\delta_{ij}$	Symbole de Kronecker
$\vartheta$	Paramètre de régularisation de Tikhonov
$\vartheta^S$	Paramètre de régularisation de Sobolev

$\vartheta_v^S$	Filtre de Sobolev spatialement dépendant
$d^m, \delta\gamma^m$	Direction de descente à l'itération $m$
$\alpha$	Pas de descente
$\text{Re}(u)$	Partie réelle de l'expression complexe $u$
$\text{Im}(u)$	Partie imaginaire de l'expression complexe $u$
$\bar{u}$	Conjugué de l'expression complexe $u$
$\mathcal{L}$	Lagrangien
$\mathcal{M}$	Maillage
$\mathcal{V}$	Espace d'approximation
$j$	Fonction de coût dépendant de la variable de contrôle $\gamma$
$\mathcal{J}$	Fonction de coût dépendant de la variable d'état $\varphi$
$\mathcal{J}^+$	Pénalisation de la fonction de coût de type Tikhonov
$j'(\gamma; \eta)$	Dérivée directionnelle de $j$ au point $\gamma$ dans la direction $\eta$
<b>E</b>	Espérance mathématique
$\mathcal{N}(0, 1)$	Loi normale centrée réduite
$e_1$	Erreur quadratique relative
$e_2$	Coefficient de corrélation

### Espaces fonctionnels

$L^\infty(\Omega)$	Espace des fonctions à valeurs réelles essentiellement bornées $L^\infty(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \text{ tel que } \exists C < +\infty :  f(\mathbf{r})  \leq C \text{ pour presque tout } \mathbf{r} \in \Omega\}$
$L^2(\Omega)$	Espace des fonctions à valeurs réelles de carré intégrable $L^2(\Omega) = \left\{f : \Omega \rightarrow \mathbb{R} \text{ tel que } \int_\Omega  f(\mathbf{r}) ^2 d\mathbf{r} < +\infty\right\}$
$\hat{L}^2(\Omega)$	Espace des fonctions à valeurs complexes de carré intégrable $\hat{L}^2(\Omega) = \left\{f : \Omega \rightarrow \mathbb{C} \text{ tel que } \int_\Omega  f(\mathbf{r}) ^2 d\mathbf{r} < +\infty\right\}$
$H^1(\Omega)$	Espace de Sobolev des fonctions à valeurs réelles

$$H^1(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \text{ tel que } f \in L^2(\Omega) \text{ et } \frac{\partial f}{\partial r_j} \in L^2(\Omega), \forall j = 1, \dots, n_\Omega \right\}$$

 $\hat{H}^1(\Omega)$ 

Espace de Sobolev des fonctions à valeurs complexes

$$\hat{H}^1(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{C} \text{ tel que } f \in \hat{L}^2(\Omega) \text{ et } \frac{\partial f}{\partial r_j} \in \hat{L}^2(\Omega), \forall j = 1, \dots, n_\Omega \right\}$$

### Opérateurs

 $(u, v)_{\mathcal{X}}$ 

Produit scalaire  $\mathcal{X}$  de  $u$  et  $v$

 $u^\top$ 

Transposé du vecteur ou de la matrice  $u$

 $\nabla$ 

Gradient

 $\nabla \cdot$ 

Divergence

 $\Delta$ 

Laplacien

### Indices

ap

*A priori*

dp

Principe de Morozov (*discrepancy principle*)

t

Cible (*target*)



# INTRODUCTION

## Contexte

Au cours des dernières décennies, on a assisté à un développement rapide de nouvelles techniques optiques de diagnostic, que ce soit dans les domaines des procédés industriels (Khairi *et al.*, 2012) ou de l'imagerie médicale (Arridge, 1999; Charette *et al.*, 2008). Ces avancées ont été rendues possibles d'une part grâce aux développements dans la compréhension théorique des transferts radiatifs et de l'interaction entre la lumière et les propriétés radiatives des milieux semi-transparents (MST). D'autre part, l'amélioration des moyens de détection de la lumière et des sources lumineuses qui irradient les milieux a été nécessaire pour la mise en pratique de ces techniques. L'une de ces avancées est la possibilité d'obtenir une image de la distribution spatiale des propriétés radiatives à partir de mesures optiques à la surface. La distribution spatiale des propriétés radiatives donne des informations sur l'hétérogénéité de ces propriétés, ce qui permet, par exemple, de remonter à l'existence de défauts, à l'état pathologique des tissus humains (diagnostics médicaux) ou encore à la physiologie d'un tissu (monitoring du taux d'oxygénation du sang).

Cette technique d'imagerie est une méthode inverse nommée la Tomographie Optique Diffuse<sup>1</sup> (TOD). Elle consiste à identifier les paramètres d'un modèle numérique (modèle direct) de transferts radiatifs à partir de mesures en frontière par minimisation de l'écart entre les mesures et la prédiction du modèle numérique.

Cette méthode présente plusieurs modalités. Dans la première, qui est l'objet de cette thèse, le milieu à caractériser est sollicité par une source extérieure monochromatique et la détection s'opère à la même longueur d'onde. Dans la seconde, appelée TOD de fluorescence, des agents fluorescents ayant la propriété, lorsqu'ils sont soumis à un rayonnement d'une certaine longueur d'onde, d'émettre à une longueur d'onde supérieure, sont injectés dans les tissus. La

---

<sup>1</sup>On trouvera également les appellations *photon migration imaging*, *photon migration tomography*, *diffuse photon density wave* ou simplement *optical tomography* dans la littérature. Nous emploierons l'expression « tomographie optique diffuse » dans ce travail et une légère distinction entre les termes DOT et OT sera considérée dans le chapitre 2.

TOD de fluorescence consiste à déterminer la distribution de ces agents à partir de la détection de leur signal, car les zones de concentration de ces agents peuvent révéler la présence de tumeurs, par exemple. Dans la troisième, appelée TOD de bioluminescence, des agents extérieurs sont également injectés dans les tissus. Cependant, à la différence de la TOD de fluorescence, ces agents ne réagissent pas à un rayonnement mais par réaction enzymatique. La localisation de ces sources d'émission traduit la présence de ces enzymes et donc de tumeurs, par exemple.

Le choix du modèle direct permet également de distinguer deux catégories : la TOD basée sur l'équation de transfert radiatif (ÉTR) et la TOD basée sur l'équation de la diffusion. L'ÉTR est le modèle de référence pour modéliser le transport des photons dans les MST. Cependant, les temps de calcul et les difficultés liées à sa résolution pour des géométries complexes font qu'elle reste peu appliquée à la TOD. L'équation de la diffusion est une approximation de l'ÉTR valide dans les milieux faiblement absorbant, fortement diffusant et sous des hypothèses de grandes échelles de temps et d'espace. Cette équation, qui est utilisée dans ce travail, est le modèle privilégié pour des applications de TOD en raison de sa facilité de résolution pour des géométries quelconques. La TOD basée sur l'équation de la diffusion consiste à déterminer les distributions spatiales des coefficients d'absorption,  $\kappa$ , et de diffusion réduit,  $\sigma$ , du milieu imagé.

Trois variantes, relatives aux types de sources et de détecteurs utilisés, sont également à distinguer : les régimes continu, fréquentiel et temporel. Les dispositifs utilisant des sources continues en intensité ont été les premiers à être développés en TOD en raison de leur faible coût et de leur simplicité. L'atténuation du signal lumineux après propagation dans le milieu est mesurée par les détecteurs. Il a été démontré qu'il n'est pas possible de prouver l'unicité des coefficients d'absorption et de diffusion réduit retrouvés par ce type de mesure (Arridge et Lionheart, 1998). L'unicité peut cependant être retrouvée en restreignant l'espace de recherche de ces coefficients (Harrach, 2009). Les systèmes fréquentiels utilisent des sources de lumière modulées en intensité à une fréquence radio comprise entre 100 MHz et 1 GHz. En plus de l'atténuation du signal moyen, ces systèmes mesurent l'atténuation de l'amplitude de modulation ainsi que le déphasage entre le signal incident et détecté. Ces informations supplémentaires permettent

de mieux déterminer les coefficients d'absorption et de diffusion réduit. Enfin, les dispositifs résolus en temps utilisent des sources de lumière impulsives ultracourtes, de l'ordre de la picoseconde voire de la femtoseconde, et mesurent la distribution temporelle de l'intensité émergente. Lorsque toute cette information est exploitée, le régime temporel est plus riche en informations que le régime fréquentiel car les données temporelles sont équivalentes, via la transformée de Fourier, aux données fréquentielles sur un très grand nombre de fréquences. Cependant, en raison du coût élevé des sources pulsées et des détecteurs associés, les dispositifs temporels sont les plus onéreux. La modalité fréquentielle, considérée comme la plus mature et la plus utilisée (Wang et Wu, 2012), a été choisie dans ce travail. Une revue des différents dispositifs expérimentaux, sources et détecteurs utilisés en TOD pourra être trouvée dans (Alonzo *et al.*, 2011; Zhang, 2014), par exemple.

Une des applications de la TOD est donc l'imagerie médicale, reposant sur le fait que la connaissance des propriétés radiatives donne des informations sur l'état physiopathologique des tissus. On peut ainsi détecter précocement les tumeurs cancéreuses, effectuer le monitoring d'enfants prématurés, contrôler le taux d'oxygénation du sang dans le cerveau, diagnostiquer suffisamment tôt l'arthrite rhumatoïde (Klose, 2001) ou encore utiliser la TOD pour la recherche médicale dans l'imagerie du petit animal. La Figure 0.1 présente par exemple un appareillage de TOD pour la détection de tumeurs dans le sein. Dans ce dispositif, les sources lumineuses sont des diodes laser émettant en continu et à 8 longueurs d'onde différentes comprises entre 638 et 915 nm, tandis que la détection des densités de photons émergentes s'effectue à l'aide de photodiodes en silicium. Le tableau 0.1 montre tout l'intérêt de la TOD par rapport aux autres principales méthodes de diagnostic que sont la tomographie informatisée par rayons-X (TI-X), l'imagerie par résonance magnétique (IRM) et la tomographie par émission de positrons (TEP) : son coût d'installation est bien inférieur à celui des trois autres méthodes et contrairement aux techniques TI-X et TEP, la TOD est une méthode non nocive puisque mettant en jeu du rayonnement proche-infrarouge (IR) de faible intensité. En revanche, sa résolution spatiale, de l'ordre du centimètre, est moins bonne que celle des autres méthodes.

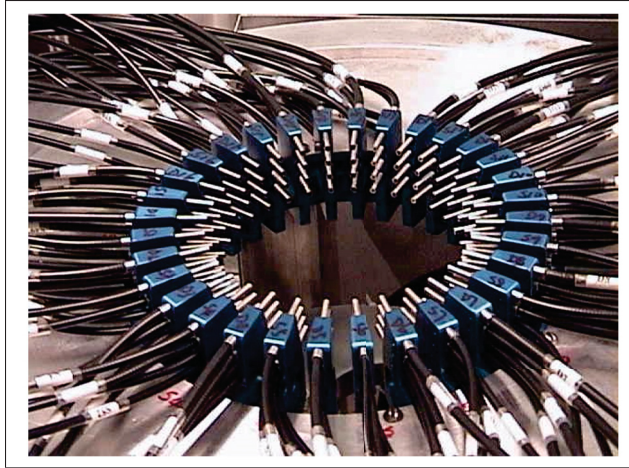


FIG. 0.1 Appareillage de TOD pour la détection de tumeurs dans le sein. Tiré de (Hajihashemi *et al.*, 2012).

TAB. 0.1 Comparaison de différentes méthodes de diagnostic. Adapté de (Charette *et al.*, 2011).

	Type de rayonnement	Traceur radioactif	Coût de l'installation	Résolution spatiale
<b>TI-X</b>	Rayons X (ionisant)	Non	Élevé	Grande ( $\approx 0.3 \mu\text{m}$ )
<b>IRM</b>	Champ magnétique et radiofréquences (non ionisant)	Non	Élevé	Grande ( $\approx 100 \mu\text{m}$ )
<b>TEP</b>	Rayons $\gamma$ (ionisant)	Oui	Élevé	Grande ( $\approx 1 \text{ mm}$ )
<b>TOD</b>	Visible/Proche IR (650 – 950 nm) (non ionisant)	Non	Faible	Faible ( $\approx 1 \text{ cm}$ )

L'imagerie du petit animal, motivé par les progrès de la génétique (Hielscher, 2005), consiste à observer sur des petits animaux tels que des rongeurs le développement de pathologies pouvant être reliées au développement de ces mêmes pathologies sur l'homme. L'intérêt de la TOD pour cette recherche porte principalement sur le fait que la non nocivité de cette méthode d'imagerie permet le suivi de la maladie sur un même animal et limite ainsi les erreurs dues à la spécificité de chaque individu.

Ainsi la tomographie optique diffuse s'associe au progrès de la médecine d'une part dans le diagnostic médical et d'autre part pour la recherche médicale. Or, la lutte contre des maladies telles que le cancer ou la maladie d'Alzheimer constitue un enjeu humain et économique pour la société. Les statistiques de la société canadienne du cancer (Comité consultatif de la Société canadienne du cancer, 2014) estiment à 191 300 nouveaux cas de cancer et 76 600 décès causés par cette maladie au Canada en 2014. Autrement dit, toutes les heures, environ 22 personnes recevront un diagnostic de cancer au Canada en 2014. Les hausses du nombre de nouveaux cas de cancer sont plus importantes chaque année et sont principalement attribuables à la croissance démographique et au vieillissement de la population. On comprend donc tout l'enjeu du développement de ce type de technologie.

D'autres applications envisageables de la TOD sont la détection de défauts volumiques dans des matériaux semi-transparents, donc le contrôle non destructif dans l'industrie des matériaux, ou encore la visualisation d'écoulements par la reconstruction de l'indice de réfraction (Bou-langer *et al.*, 2011). La caractérisation de matériaux semi-transparents comme les composites thermoplastiques à renforts fibreux peut faire également partie des applications de la TOD. Selon la nature du procédé utilisé, l'inhomogénéité en répartition des fibres peut induire des variations spatiales des propriétés radiatives accessibles aux résolutions offertes par la TOD. La détermination des propriétés radiatives de ces composites dans l'IR est d'une importance cruciale car de nombreux procédés industriels utilisent des sources IR (thermoformage ou pultrusion) pour la fabrication de ces matériaux. Une connaissance fine des propriétés radiatives est donc nécessaire pour l'optimisation de ces procédés. Dans la plupart des cas, les propriétés radiatives volumiques de ces matériaux sont homogènes dans le milieu de sorte que la détermination de leurs distributions spatiales n'est pas nécessaire. Il s'agit plutôt de déterminer les coefficients d'absorption et de diffusion homogénéisés, ainsi que les coefficients de la fonction de phase, réduisant le nombre de paramètres à identifier à quelques unités contre des centaines ou milliers de paramètres lors d'une reconstruction spatiale. D'autres algorithmes, inapplicables pour la reconstructions de propriétés variant dans l'espace, sont ainsi généralement utilisés (Baillis et Sacadura, 2000). Toutefois, les matériaux peuvent présenter par exemple des

gradients de répartition en fibres, auquel cas les propriétés radiatives dépendront de l'espace et l'utilisation d'algorithmes de reconstruction développés dans le cadre de la TOD sera requise.

## Problématique

Le rayonnement rouge/proche IR est utilisé en TOD car les tissus biologiques sont peu absorbants à ces longueurs d'onde (voir Figure 0.2) et donc les photons peuvent y pénétrer profondément voire les traverser complètement, et permettre ainsi un sondage des tissus en profondeur. Cette gamme de fréquence est appelée fenêtre thérapeutique, comprise entre 650 et 950nm. En contrepartie, le rayonnement est fortement diffusant dans ces tissus à ces longueurs d'onde. Les photons ne suivant pas des trajectoires rectilignes, le rayonnement émergent est donc mal adapté à l'imagerie directe comme il l'est en TI-X.

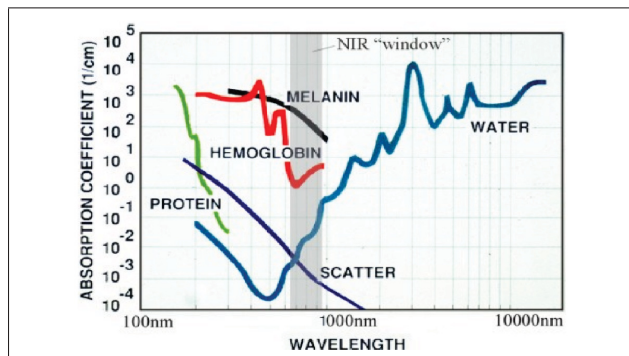


FIG. 0.2 Représentation de la fenêtre thérapeutique : coefficient d'absorption des principaux composants absorbants dans les tissus biologiques en fonction de la longueur d'onde. Extrait de (Hillman, 2002).

Malgré cela, la résolution d'un problème inverse permet d'exploiter tout de même l'information contenue dans ce rayonnement émergent. La Figure 0.3 présente un schéma du principe de la TOD. On observe sur ce schéma que la TOD nécessite la résolution d'un triple problème. Premièrement, il faut résoudre un problème d'instrumentation (en haut à gauche) : exposant le milieu à caractériser à des sources lumineuses, il s'agit d'obtenir des mesures précises et ap-

portant le maximum d'information afin d'effectuer la reconstruction des paramètres  $\gamma$  le plus aisément possible. Ensuite, il faut résoudre le problème direct (en haut à droite) : supposant le vecteur des paramètres à retrouver  $\gamma$  connu, il s'agit de décrire la propagation de la lumière dans le milieu et d'en extraire les mesures prédites aux capteurs. À partir des mesures expérimentales et des prédictions numériques, il est possible de définir une fonction de coût  $j(\gamma) \geq 0$  mesurant l'écart quadratique entre ces deux grandeurs (centre). Remarquant que cette fonction décroît lorsque les prédictions se rapprochent des mesures, et donc que les valeurs du vecteur  $\gamma$  tendent vers les valeurs réelles des paramètres à retrouver, la minimisation de la fonction de coût constitue le problème inverse (en bas à droite) : il s'agit de mettre en place un algorithme de minimisation qui, partant d'un vecteur initial  $\gamma_0$ , fait converger ce vecteur vers celui des paramètres à retrouver, à l'aide d'un processus itératif. Quand un critère de convergence est satisfait, l'algorithme s'arrête et une estimation des paramètres est obtenue (en bas à gauche).

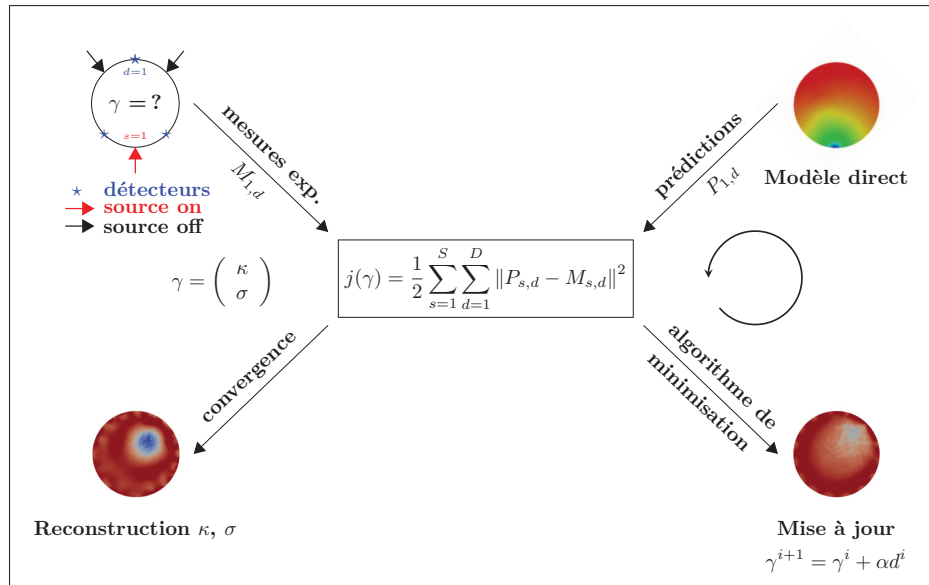


FIG. 0.3 Schématisation du principe général du problème de tomographie optique.

De nos jours, des problèmes restent à surmonter pour permettre d'obtenir des algorithmes de minimisation rapides et précis qui transforment les mesures optiques en image de la distribution spatiale des propriétés radiatives. Il s'agit entre autres de développer des modèles numériques



de propagation de la lumière dans les MST capables de prendre en compte les géométries complexes avec des temps de calcul raisonnables et une précision acceptable. Aussi, la netteté des images est à améliorer car le problème inverse est mal posé. En effet, d'une part, celui-ci est sensible au bruit inhérent aux mesures expérimentales. D'autre part, il est sous-déterminé car le nombre de mesures expérimentales est généralement bien inférieur au nombre de paramètres à estimer. Le développement de méthodes de régularisation adaptées est au cœur de ce problème.

## **Objectifs**

Le but de ce projet de doctorat est le développement d'une méthode inverse robuste pour la résolution du problème de TOD dans le domaine fréquentiel. Pour répondre à cet objectif, trois thèmes ont été investigués qui sont respectivement :

- a. Développement d'un code numérique de résolution du modèle de l'approximation diffuse bidimensionnel et tridimensionnel dans le régime fréquentiel par la méthode des éléments finis ;
- b. Développement d'algorithmes de minimisation robustes et performants ;
- c. Développement de méthodes de régularisation adaptées au caractère mal-posé du problème inverse.

## **Contenu de la thèse**

Huit chapitres et quatre annexes viennent s'ajouter à l'introduction. Le chapitre 1 présente la revue bibliographique. Les chapitres 2, 3, 4 et 5 présentent les articles publiés ou soumis à des journaux scientifiques avec comité de lecture. Le chapitre 6 présente des résultats complémentaires obtenus au cours de cette thèse et non présentés dans les articles. Le chapitre 7 fournit une discussion des résultats. Le chapitre 8 présente la conclusion et les perspectives. L'annexe 1 est une annexe du chapitre 2. L'annexe 2 est un article de conférence présentant des résultats utilisés dans les chapitres 2, 4 et 5. L'annexe 3 propose une validation de la résolution du modèle direct. Enfin, l'annexe 4 donne la liste des publications en lien avec la thèse.



## CHAPITRE 1

### REVUE DE LITTÉRATURE

Le présent chapitre vise à fournir les quelques notions de base du transfert radiatif en champ lointain jugées indispensables à la bonne compréhension physique des problèmes inverses résolus dans la suite de ce document. Dans un premier temps, les grandeurs photométriques et caractéristiques du transport de la lumière dans les milieux semi-transparents sont introduites. L'équation de transfert radiatif (ÉTR) est ensuite dérivée selon l'approche dite phénoménologique. Enfin, le modèle direct de propagation de la lumière considéré dans cette thèse, dit de l'approximation diffuse, est rigoureusement obtenu à partir de l'ÉTR.

La plupart des informations qui suivent sont issues de livres de référence tels que (Modest, 2013), (Mahan, 2002) ou (Howell *et al.*, 2011). Cette thèse considère l'approche corpusculaire de la lumière dans laquelle le transport de la lumière est assimilé à un transport de photons interagissant les uns les autres de façon indépendante dans un milieu de propagation supposé fixe mais dont les éléments sont distribués aléatoirement dans l'espace (Sacadura, 2011). Le photon est modélisé par une particule située à un point de l'espace, ayant une direction  $\mathbf{s}$ , se déplaçant à la vitesse  $c$  et possédant l'énergie  $h\nu$  où  $\nu$  est sa fréquence et  $h = 6,626 \times 10^{-34} \text{Js}$  est la constante de Planck. Ainsi, les variables utilisées pour décrire la propagation des photons dépendront notamment de la position dans l'espace et de la direction de propagation. Dans ce document,  $\Omega \subset \mathbb{R}^{n_\Omega}$  représente le domaine d'étude,  $\mathbf{r} = (r_1, \dots, r_{n_\Omega})$  le vecteur position en coordonnées cartésiennes et  $\mathbf{s} = (s_1, s_2)$  la direction de propagation, en coordonnées sphériques. La grandeur appropriée pour décrire les transferts radiatifs à l'échelle mésoscopique est la luminance, quantité locale et directionnelle.

## 1.1 Grandeurs photométriques

### 1.1.1 Luminance

La luminance  $L$  est définie de telle sorte que le flux d'énergie radiative monochromatique  $dP$  traversant la surface élémentaire  $dA$  centré au point  $\mathbf{r}$ , dans l'angle solide élémentaire  $ds$  centré sur la direction  $\mathbf{s}$ , dans l'intervalle de fréquences  $[\nu, \nu + d\nu]$ , au temps  $t$ , vaut :

$$dP(\mathbf{r}, \mathbf{s}, \nu, t) = L(\mathbf{r}, \mathbf{s}, \nu, t) \mathbf{s} \cdot \mathbf{n} ds d\nu dA \quad (1.1)$$

où  $\mathbf{n}$  est le vecteur normal unitaire à  $dA$ . La luminance s'interprète comme une quantité d'énergie de fréquence  $\nu$  se propageant selon une direction et un angle solide donnés, au point  $\mathbf{r}$  et au temps  $t$ . Son unité s'exprime en  $\text{W m}^{-2} \text{sr}^{-1} \text{Hz}^{-1}$ . La luminance totale, d'unité le  $\text{W m}^{-2} \text{sr}^{-1}$ , est, quant à elle, définie comme l'intégrale de la luminance sur l'ensemble des fréquences. Comme cette thèse limite son attention à une fréquence fixée, la variable  $\nu$  est implicite et la luminance monochromatique  $L_\nu(\mathbf{r}, \mathbf{s}, t)$  est utilisée (l'indice  $\nu$  de  $L_\nu$  sera toutefois omis). Dans la suite de ce chapitre, les grandeurs calculées à partir de la luminance seront définies par rapport à la luminance totale.

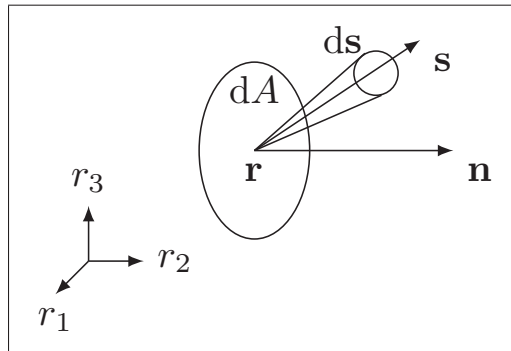


FIG. 1.1 Géométrie utilisée pour la définition de la luminance.

### 1.1.2 Vecteur flux radiatif

Le vecteur flux radiatif  $\mathbf{J}$ , d'unité le  $\text{W m}^{-2}$ , est défini par :

$$\mathbf{J}(\mathbf{r}, t) = \int_{S^{\Omega-1}} L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} d\mathbf{s} \quad (1.2)$$

Le flux radiatif total émis par l'élément de surface  $dA$  est alors :

$$dP(\mathbf{r}, t) = \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n} dA \quad (1.3)$$

### 1.1.3 Densité volumique d'énergie

La densité volumique d'énergie  $U(\mathbf{r}, t)$  est définie telle que :

$$U(\mathbf{r}, t) = \frac{1}{c} \int_{S^{\Omega-1}} L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \quad (1.4)$$

Son unité est le  $\text{J m}^{-3}$  et cette grandeur s'interprète comme la quantité d'énergie radiative totale par unité de volume au point  $\mathbf{r}$  et au temps  $t$ .

### 1.1.4 Densité de photons

Enfin, la densité de photons  $\varphi$ , d'unité le  $\text{W m}^{-2}$ , est définie telle que :

$$\varphi(\mathbf{r}, t) = c \times U(\mathbf{r}, t) \quad (1.5)$$

Cette dernière grandeur, ne dépendant pas de la direction, sera utilisée par la suite pour obtenir une équation pouvant décrire les transferts radiatifs dans les milieux semi-transparents.

## 1.2 Grandeurs caractéristiques du transport de la lumière

### 1.2.1 Indice de réfraction

L'indice de réfraction  $n$  d'un milieu caractérise la vitesse de propagation  $c$  de la lumière dans celui-ci par la relation :

$$n = \frac{c_0}{c} \quad (1.6)$$

où  $c_0 = 299\,792\,458\text{ m s}^{-1}$  est la vitesse de la lumière dans le vide. À l'échelle macroscopique, l'indice de réfraction des tissus biologiques varie peu et est de l'ordre de 1,40. Il dépend en fait essentiellement de la teneur en eau du tissu, 1,33 étant l'indice de réfraction de l'eau et 1,50 celui estimé d'un tissu déshydraté (Torregrossa, 2003).

### 1.2.2 Absorption

Considérons un rayon lumineux se propageant dans un milieu participant dans la direction  $\mathbf{s}$  selon l'axe  $Or_1$ . L'énergie radiative portée par ce rayon va subir le long de son parcours une atténuation par absorption de photons. On observe que la quantité absolue d'énergie absorbée est proportionnelle à la quantité d'énergie incidente et à la distance parcourue (Modest, 2013). On peut alors écrire :

$$dL(r_1, \mathbf{s}, t) = -\kappa L(r_1, \mathbf{s}, t) dr_1 \quad (1.7)$$

Le coefficient de proportionnalité,  $\kappa$ , est appelé coefficient d'absorption. Sa valeur dans les tissus biologiques varie entre 0,01 et 0,5 cm<sup>-1</sup> (Hielscher *et al.*, 2011; Bashkatov *et al.*, 2011). L'intégration de l'équation (1.7) le long d'un chemin de longueur  $l$  fournit la loi d'atténuation exponentielle du rayonnement connue sous le nom de loi de Beer-Lambert :

$$L(l, \mathbf{s}, t) = L(0, \mathbf{s}, t) e^{-\kappa l} \quad (1.8)$$

### 1.2.3 Émission

Tout milieu à une température supérieure au zéro absolu émet du rayonnement électromagnétique. La quantité d'énergie radiative émise par un élément de volume est proportionnelle au volume de cet élément. Ce faisant :

$$dL(r_1, \mathbf{s}, t) = E(r_1, \mathbf{s}, t) dr_1 \quad (1.9)$$

où  $E$  est une fonction caractérisant le mécanisme d'émission. À l'équilibre thermodynamique local, c'est-à-dire lorsque la matière est localement proche de l'équilibre thermique, mécanique et osmotique, cette énergie est liée à la luminance telle que (Chandrasekhar, 1960) :

$$E(r_1, \mathbf{s}, t) = \kappa L^0(r_1, t, T) \quad (1.10)$$

où  $T$  est la température locale et  $L^0 = \frac{2h\nu^3}{c^2 e^{\frac{h\nu}{k_\beta T}} - 1}$  est la luminance de Planck,  $k_\beta$  étant la constante de Boltzmann.

### 1.2.4 Diffusion de la lumière

Le long de son trajet dans la direction  $\mathbf{s}$ , une partie de l'énergie incidente est diffusée dans d'autres directions, contribuant à sa diminution. Comme pour l'absorption, la quantité d'énergie diffusée est proportionnelle à la quantité d'énergie incidente et à la distance parcourue. Ainsi :

$$dL(r_1, \mathbf{s}, t) = -\sigma_s L(r_1, \mathbf{s}, t) dr_1 \quad (1.11)$$

où  $\sigma_s$  est appelée coefficient de diffusion. Sa valeur varie entre 20 et 200 cm<sup>-1</sup> dans les tissus biologiques (Cheong *et al.*, 1990; Bashkatov *et al.*, 2011). La suite de ce travail suppose que la diffusion est uniquement élastique, de sorte que pour une fréquence fixée il n'y a pas de perte d'énergie à chaque évènement de diffusion par changement de fréquence du rayonnement.

Le coefficient d'extinction est défini tel que  $\beta = \kappa + \sigma_s$ . L'inverse est appelé longueur d'extinction et caractérise la distance parcourue par un photon avant de subir une interaction de type absorption ou diffusion. On définit aussi l'épaisseur optique :

$$\tau(l) = \int_0^l \beta(r_1) dr_1 \quad (1.12)$$

Pour un milieu de diamètre  $l$ , si  $\tau(l) \ll 1$ , le milieu est dit optiquement mince et si  $\tau(l) \gg 1$ , il est dit optiquement épais.

### 1.2.5 Fonction de phase

Si le phénomène de diffusion peut conduire à une diminution de l'intensité du rayon lumineux le long de son trajet dans la direction  $\mathbf{s}$ , il peut aussi conduire à son augmentation, lorsqu'un photon incident selon une direction  $\mathbf{s}'$  est diffusé dans la direction  $\mathbf{s}$ . La fonction de phase  $\Psi$  tente de quantifier ce phénomène. Elle constitue une densité de probabilité sur l'ensemble des directions et peut-être interprétée physiquement comme le rapport de la luminance diffusée dans une direction sur la luminance diffusée dans cette direction dans le cas d'une diffusion isotrope (Rousse, 1994). En particulier, elle doit vérifier

$$\int_{S^{\Omega-1}} \Psi(\mathbf{s}', \mathbf{s}) d\mathbf{s}' = 1 \quad (1.13)$$

On suppose très souvent que le milieu n'a pas de direction préférentielle et que les particules diffusantes sont composées d'un matériau homogène, isotrope et présentant une symétrie sphérique parfaite, et qu'ainsi la fonction de phase ne dépend en fait que du cosinus de l'angle relatif entre les deux directions  $\mathbf{s}$  et  $\mathbf{s}'$  (Moura, 1998). En notant  $\theta$  cet angle relatif, par définition  $\cos(\theta) = \mathbf{s} \cdot \mathbf{s}'$  le produit scalaire usuel de  $\mathbf{s}$  et  $\mathbf{s}'$ . De manière générale, la fonction de phase peut être représentée par un développement en série de polynômes de Legendre :

$$\Psi(\cos(\theta)) = 1 + \sum_{n=1}^{\infty} A_n P_n(\cos(\theta)) \quad (1.14)$$

où la tabulation des coefficients  $A_n$  est effectuée dans (Chu *et al.*, 1957). Un paramètre important associé à la fonction de phase est le facteur d'asymétrie  $g \in [-1, 1]$ . Ce paramètre caractérise la tendance de la diffusion : pour  $g = 0$ , la diffusion est isotrope, pour  $g < 0$ , la diffusion se fait préférentiellement vers l'arrière et pour  $g > 0$ , la diffusion se fait préférentiellement vers l'avant (les cas limites  $g = 1$  et  $g = -1$  correspondant à une diffusion totale vers l'avant et une rétrodiffusion totale vers l'arrière, respectivement). Il est défini par :

$$g = \int_{S^{n_{\Omega}-1}} \mathbf{s} \cdot \mathbf{s}' \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s} \quad (1.15)$$

La fonction de phase couramment utilisée pour la propagation de la lumière dans les tissus biologiques est la fonction de phase d'Henye-Greenstein (Henyey et Greenstein, 1941), initialement développée par ces auteurs pour caractériser la diffusion dans les nuages de poussière interstellaire. Cette fonction possède la particularité de dépendre explicitement du facteur d'asymétrie  $g$ . Elle est définie par :

$$\Psi(\cos(\theta)) = \begin{cases} \frac{1 - g^2}{2\pi(1 + g^2 - 2g\cos(\theta))} & \text{cas bidimensionnel} \\ \frac{1 - g^2}{4\pi(1 + g^2 - 2g\cos(\theta))^{3/2}} & \text{cas tridimensionnel} \end{cases} \quad (1.16)$$

Les valeurs typiques du facteur d'asymétrie dans les tissus biologiques se situent dans l'intervalle  $[0, 0.98]$  (Cheong *et al.*, 1990; Bashkatov *et al.*, 2011).

Finalement, le gain d'énergie radiative provenant de la diffusion dans toutes les directions s'écrit :

$$dL(r_1, \mathbf{s}, t) = \sigma_s \left( \int_{S^{n_{\Omega}-1}} L(r_1, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right) dr_1 \quad (1.17)$$

Une revue assez complète de différentes fonctions de phase est fournie dans la thèse de (Moura, 1998).

### 1.3 L'équation de transfert radiatif

#### 1.3.1 Dérivation de l'équation de transfert radiatif selon l'approche phénoménologique

L'équation de transfert radiatif (ÉTR) est obtenue par une approche dite phénoménologique, c'est-à-dire en effectuant le bilan d'énergie radiative du rayon lumineux le long de la direction  $\mathbf{s}$  (Chandrasekhar, 1960). D'après ce qui précède, (Modest, 2013) :

$$L(r_1 + dr_1, \mathbf{s}, t + dt) - L(r_1, \mathbf{s}, t) = \kappa L^0(r_1, t, T) dr_1 - \kappa L(r_1, \mathbf{s}, t) dr_1 - \sigma_s L(r_1, \mathbf{s}, t) dr_1 + \sigma_s \left( \int_{S^{n_{\Omega}-1}} L(r_1, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right) dr_1 \quad (1.18)$$

Le premier terme de cette équation s'approxime par un développement de Taylor au premier ordre :

$$L(r_1 + dr_1, \mathbf{s}, t + dt) = L(r_1, \mathbf{s}, t) + \frac{\partial L}{\partial t} dt + \frac{\partial L}{\partial r_1} dr_1 \quad (1.19)$$

Or la vitesse de propagation du rayonnement à la vitesse  $c$  implique  $\frac{dr_1}{dt} = c$ . Ainsi on obtient :

$$\frac{1}{c} \frac{\partial L}{\partial t} + \frac{\partial L}{\partial r_1} = \kappa L^0(r_1, t, T) - (\kappa + \sigma_s) L(r_1, \mathbf{s}, t) + \sigma_s \left( \int_{S^{n_{\Omega}-1}} L(r_1, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right) \quad (1.20)$$

Tous les développements précédents supposaient que la direction  $\mathbf{s}$  était selon l'axe  $Or_1$  mais restent valables pour toutes les directions. Considérant l'ensemble des directions, l'ÉTR est finalement obtenue :

$$\left( \frac{1}{c} \frac{\partial}{\partial t} + \mathbf{s} \cdot \nabla \right) L(\mathbf{r}, \mathbf{s}, t) = -(\kappa + \sigma_s) L(\mathbf{r}, \mathbf{s}, t) + \sigma_s \left( \int_{S^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right) + S(\mathbf{r}, \mathbf{s}, t) \quad (1.21)$$

où l'émission thermique est incluse dans le terme source  $S(\mathbf{r}, \mathbf{s}, t)$ . Fait à noter, il est également possible d'établir l'ÉTR à partir des équations de Maxwell (Pierrat, 2007).



### 1.3.2 Conditions de bord

Afin de pouvoir résoudre l'ÉTR, il faut spécifier les valeurs des luminances entrantes au bord du domaine. Deux modélisations des conditions de bord sont typiquement employées dans les travaux de TO basés sur l'inversion de l'ÉTR. La première approche (*e.g.* (Balima *et al.*, 2011b; Tarvainen *et al.*, 2011)) considère la frontière transparente, *i.e.* tout rayonnement incident est transmis quelle que soit sa direction. Les seuls photons entrants dans le milieu sont ainsi ceux provenant des sources, et les conditions de bord sont données par :

$$L(\mathbf{r}, \mathbf{s}, t) = \begin{cases} \xi(\mathbf{r}, \mathbf{s}, t) & \mathbf{s} \cdot \mathbf{n} < 0, \mathbf{r} \in \partial\Omega_s \\ 0 & \mathbf{s} \cdot \mathbf{n} < 0, \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.22)$$

où  $\partial\Omega_s$  et  $\xi$  représentent l'emplacement et les valeurs des luminances des sources extérieures, respectivement.

La seconde approche (*e.g.* (Klose *et al.*, 2002; Kim et Hielscher, 2009)) considère la frontière semi-transparente, *i.e.* la différence d'indices de réfraction entre le milieu ambiant (air) et le milieu à caractériser est considérée. Cette différence impose la prise en compte des réflexions et réfractions au bord du domaine. Dans ces travaux, les auteurs se placent dans le cadre de l'optique géométrique, où seules les réflexions spéculaires sont considérées.

Dans cette seconde approche, deux types de luminances entrantes sont à distinguer. D'une part, la composante réfractée des sources de lumière extérieures constituent des luminances entrantes. D'autre part, lorsqu'un rayon provenant de l'intérieur du domaine d'étude rencontre le bord du domaine, une partie de ce rayonnement s'échappera en étant réfractée, et une autre partie sera réfléchi à l'intérieur du domaine. Les conditions de bord sont alors données par (Gao et Zhao, 2009) :

$$L(\mathbf{r}, \mathbf{s}, t) = \begin{cases} R_1(\mathbf{r}, \mathbf{s}')L(\mathbf{r}, \mathbf{s}', t) + (1 - R_2(\mathbf{r}, \mathbf{s}^*))\xi(\mathbf{r}, \mathbf{s}^*, t) & \mathbf{s} \cdot \mathbf{n} < 0, \mathbf{r} \in \partial\Omega_s \\ R_1(\mathbf{r}, \mathbf{s}')L(\mathbf{r}, \mathbf{s}', t) & \mathbf{s} \cdot \mathbf{n} < 0, \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.23)$$

où  $\mathbf{s}'$ , symétrique de  $\mathbf{s}$  par rapport à la normale  $\mathbf{n}$ , est une direction de luminance sortante réfléchie dans la direction  $\mathbf{s}$ ,  $\mathbf{s}^*$  est une direction de luminance entrante réfractée dans la direction  $\mathbf{s}$ ,  $R_1$  et  $R_2$  sont les facteurs de réflexion de Fresnel et  $\xi$  représente les sources lumineuses extérieures.

Pour le calcul du facteur de réflexion  $R_1$ , (Jackson et Fox, 1999) proposent :

$$R_1(\theta) = \begin{cases} \left( \frac{n - n_a}{n + n_a} \right)^2, & \text{si } \theta = 0 \\ 1, & \text{si } |n \sin(\theta)/n_a| > 1 \\ \frac{1}{2} \left( \frac{\sin^2(\theta' - \theta)}{\sin^2(\theta' + \theta)} + \frac{\tan^2(\theta' - \theta)}{\tan^2(\theta' + \theta)} \right), & \text{sinon} \end{cases} \quad (1.24)$$

où  $\cos(\theta) = \mathbf{s}' \cdot \mathbf{n}_1$  et  $\theta'$  vérifie la loi de Snell (Jackson et Fox, 1999), soit  $n_a \sin(\theta') = n \sin(\theta)$  pour  $|n \sin(\theta)/n_a| \leq 1$  (voir Figure 1.2).

Le calcul du facteur de réflexion  $R_2$  s'obtient de manière analogue en considérant  $\cos(\theta) = \mathbf{s}^* \cdot \mathbf{n}_2$  (voir Figure 1.3) et un changement dans la seconde égalité de (1.24), à savoir  $R_2 = 1$  si  $|n_a \sin(\theta)/n| > 1$  (ce qui n'est jamais le cas en pratique puisque les milieux considérés ont toujours un indice de réfraction plus élevé que celui que l'air).

L'ÉTR permet donc de décrire la propagation du rayonnement à l'échelle mésoscopique et demeure le modèle direct de référence pour des applications en TO en raison de son large domaine de validité. Elle est cependant particulièrement difficile à résoudre car c'est une équation intégral-différentielle en dimension 6 et ses solutions analytiques ne sont disponibles que dans des cas simples. Le recours aux méthodes numériques apparaît donc indispensable pour approcher ses solutions dès lors, par exemple, que la géométrie du domaine d'étude est complexe. La littérature est très vaste concernant les méthodes numériques de résolution de l'ÉTR. À titre d'exemple, il est possible de mentionner les méthodes populaires de Monte Carlo (Hammersley et Handscomb, 1964; Howell, 1968, 1998), de lancer de rayons (Christiansen, 1883; Li *et al.*, 1997; Tan *et al.*, 2004), des harmoniques sphériques (Jeans, 1917; Modest, 2013), des volumes finis (Raithby et Chui, 1990; Chai *et al.*, 1994), des éléments finis (Razzaque *et al.*,

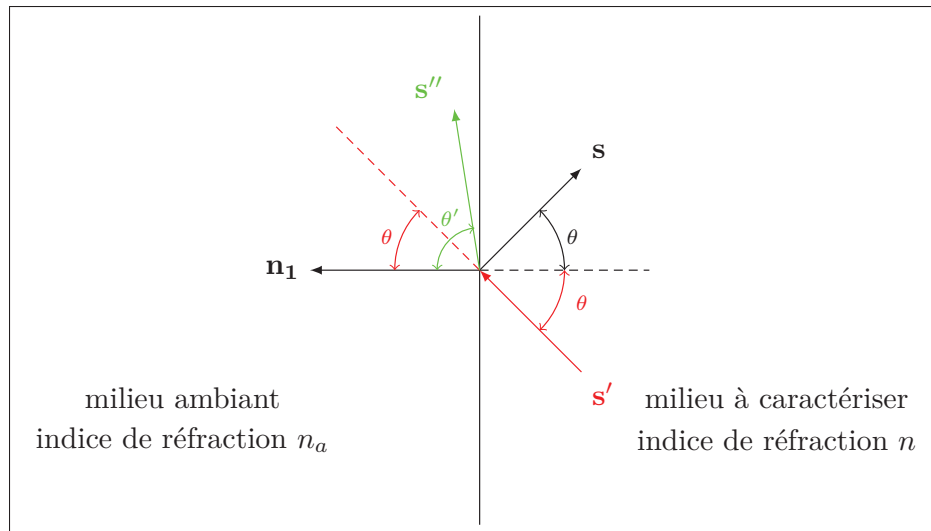


FIG. 1.2 Relation entre les rayonnements incident  $s'$ , réfléchi  $s$  et réfracté  $s''$  pour un rayonnement incident quittant le milieu.

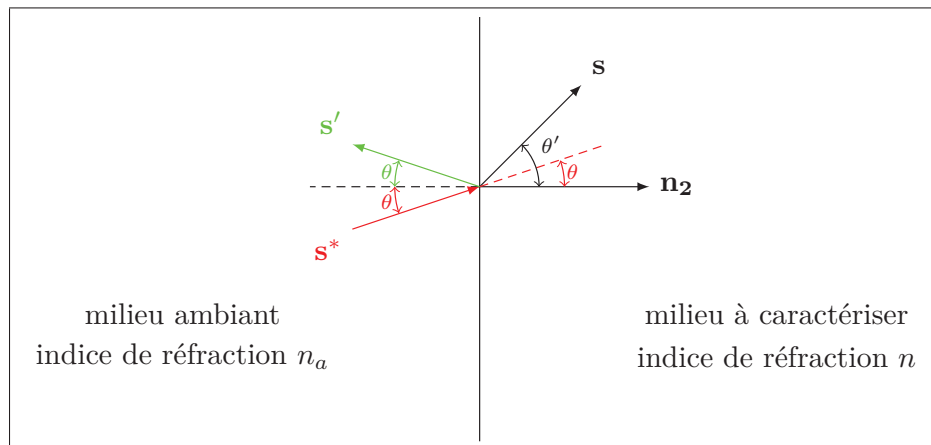


FIG. 1.3 Relation entre les rayonnements incident  $s^*$ , réfléchi  $s'$  et réfracté  $s$  pour un rayonnement incident entrant dans le milieu.

1983, 1984), des ordonnées discrètes (Chandrasekhar, 1960; Fiveland, 1984, 1988; Modest, 2013), ou encore des éléments finis basés sur des volumes de contrôle (Rousse, 2000; Rousse *et al.*, 2000).

Il convient cependant de noter que, peu importe la méthode utilisée, le calcul numérique des solutions de l'ÉTR reste relativement coûteux en temps de calcul, notamment du point de vue de la méthode inverse. En effet, le problème inverse de TO étant non-linéaire, sa résolution se

fait de manière itérative ce qui implique de faire appel à la résolution du modèle direct un très grand nombre de fois.

Dans cette thèse, l'équation de la diffusion est utilisée en tant que modèle de propagation du rayonnement. Il s'agit du modèle de l'approximation de diffusion ou approximation diffuse (*diffuse approximation* en anglais). Ce modèle direct étant beaucoup plus rapide à résoudre que l'ÉTR, ce choix a permis de concentrer cette étude davantage sur les outils de la méthode inverse.

#### 1.4 De l'équation de transfert radiatif à l'approximation diffuse

Cette section décrit le passage de l'ÉTR vers l'approximation de diffusion (AD) à partir de l'approximation P1.

##### 1.4.1 Équation sur la densité de photons

Afin d'obtenir une équation sur la densité de photons, les différents termes de l'ÉTR sont intégrés sur les directions de propagation  $\mathbf{s}$ .

- Terme de transport :

$$\int_{S^{n_{\Omega}-1}} \frac{1}{c} \frac{\partial L}{\partial t}(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = \frac{1}{c} \frac{\partial}{\partial t} \left[ \int_{S^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \right] \quad (1.25)$$

$$= \frac{1}{c} \frac{\partial \phi}{\partial t}(\mathbf{r}, t) \quad (1.26)$$

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} \cdot \nabla L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = \int_{S^{n_{\Omega}-1}} \nabla \cdot [\mathbf{s} L(\mathbf{r}, \mathbf{s}, t)] d\mathbf{s} \quad (1.27)$$

$$= \nabla \cdot \left[ \int_{S^{n_{\Omega}-1}} \mathbf{s} L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \right] \quad (1.28)$$

$$= \nabla \cdot \mathbf{J}(\mathbf{r}, t) \quad (1.29)$$

- Terme puits :

$$\int_{S^{n_{\Omega}-1}} -\beta L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = -\beta \varphi(\mathbf{r}, t) \quad (1.30)$$

- Terme source par diffusion :

$$\int_{S^{n_{\Omega}-1}} \sigma_s \left[ \int_{S^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right] d\mathbf{s} = \sigma_s \int_{S^{n_{\Omega}-1}} \Psi(\mathbf{s} \cdot \mathbf{s}') \left[ \int_{S^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}', t) d\mathbf{s}' \right] d\mathbf{s} \quad (1.31)$$

$$= \sigma_s \varphi(\mathbf{r}, t) \int_{S^{n_{\Omega}-1}} \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s} \quad (1.32)$$

D'où, d'après la relation (1.13) :

$$\int_{S^{n_{\Omega}-1}} \sigma_s \left[ \int_{S^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right] d\mathbf{s} = \sigma_s \varphi(\mathbf{r}, t) \quad (1.33)$$

- Terme source :

$$S_{\varphi}(\mathbf{r}, t) = \int_{S^{n_{\Omega}-1}} S(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \quad (1.34)$$

L'équation suivante est finalement obtenue :

$$\frac{1}{c} \frac{\partial \varphi}{\partial t}(\mathbf{r}, t) + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = -\kappa \varphi(\mathbf{r}, t) + S_{\varphi}(\mathbf{r}, t) \quad (1.35)$$

### 1.4.2 Équation sur le vecteur flux radiatif

Une équation sur le vecteur flux radiatif est obtenue en multipliant l'ÉTR par  $\mathbf{s}$  puis en intégrant sur les directions de propagation  $\mathbf{s}$ .

- Terme de transport :

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} \frac{1}{c} \frac{\partial L}{\partial t}(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = \frac{1}{c} \frac{\partial}{\partial t} \left[ \int_{S^{n_{\Omega}-1}} \mathbf{s} L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \right] \quad (1.36)$$

$$= \frac{1}{c} \frac{\partial \mathbf{J}}{\partial t}(\mathbf{r}, t) \quad (1.37)$$

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} (\mathbf{s} \cdot \nabla L(\mathbf{r}, \mathbf{s}, t)) d\mathbf{s} = \int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} \nabla L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \quad (1.38)$$

Afin d'exprimer cette intégrale à l'aide des descripteurs  $\varphi$  et  $\mathbf{J}$ , il faut faire une approximation sur la distribution angulaire de la luminance. L'approximation P1 consiste à modéliser la luminance  $L$  par son développement en harmoniques sphériques à l'ordre 1, donné par (Kaipio et Somersalo, 2005, p. 213) :

$$L(\mathbf{r}, \mathbf{s}, t) = \frac{1}{|S^{n_{\Omega}-1}|} \varphi(\mathbf{r}, t) + \frac{n_{\Omega}}{|S^{n_{\Omega}-1}|} \mathbf{s} \cdot \mathbf{J}(\mathbf{r}, t) \quad (1.39)$$

Alors :

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} \nabla L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = \frac{1}{|S^{n_{\Omega}-1}|} \left( \int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} \nabla \varphi(\mathbf{r}, t) d\mathbf{s} + n_{\Omega} \int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} \nabla [\mathbf{s} \cdot \mathbf{J}(\mathbf{r}, t)] d\mathbf{s} \right) \quad (1.40)$$

Remarquant que la fonction  $f(\mathbf{s}) = \mathbf{s} \mathbf{s}^{\top} \nabla [\mathbf{s} \cdot \mathbf{J}(\mathbf{r}, t)]$  vérifie  $f(\mathbf{s}) = -f(-\mathbf{s})$ , il est possible de déduire que la dernière intégrale de (1.40) est nulle. D'où

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} \nabla L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = \frac{1}{|S^{n_{\Omega}-1}|} \left[ \int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} d\mathbf{s} \right] \nabla \varphi(\mathbf{r}, t) \quad (1.41)$$

$$= \frac{\nabla \varphi(\mathbf{r}, t)}{n_{\Omega}} \quad (1.42)$$

où, en notant  $\text{Id}$  la matrice identité

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} \mathbf{s}^{\top} d\mathbf{s} = \frac{|S^{n_{\Omega}-1}|}{n_{\Omega}} \text{Id} \quad (1.43)$$

- Terme puits :

$$\int_{S^{n_{\Omega}-1}} -\beta \mathbf{s} L(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} = -\beta \mathbf{J}(\mathbf{r}, t) \quad (1.44)$$

- Terme source par diffusion :

$$\int_{S^{n_{\Omega}-1}} \mathbf{s} \sigma_s \left[ \int_{S^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}', t) \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}' \right] d\mathbf{s} = \sigma_s \int_{S^{n_{\Omega}-1}} \left[ \int_{S^{n_{\Omega}-1}} \mathbf{s} \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s} \right] L(\mathbf{r}, \mathbf{s}', t) d\mathbf{s}' \quad (1.45)$$

$$= \sigma_s \int_{S^{n_{\Omega}-1}} g \mathbf{s}' L(\mathbf{r}, \mathbf{s}', t) d\mathbf{s}' \quad (1.46)$$

$$= \sigma_s g \mathbf{J}(\mathbf{r}, t) \quad (1.47)$$

Dans ce développement, la définition du facteur d'asymétrie a été utilisée pour le calcul du vecteur  $\mathbf{V} = \int_{S^{n_{\Omega}-1}} \mathbf{s} \Psi(\mathbf{s} \cdot \mathbf{s}') d\mathbf{s}$  (voir section 1.2.5). L'égalité  $g = \mathbf{V} \cdot \mathbf{s}'$  étant vérifiée pour tout  $\mathbf{s}' \in S^{n_{\Omega}-1}$ , alors  $\mathbf{V} = g \mathbf{s}'$ .

- Terme source :

$$S_{\mathbf{J}}(\mathbf{r}, t) = \int_{S^{n_{\Omega}-1}} \mathbf{s} S(\mathbf{r}, \mathbf{s}, t) d\mathbf{s} \quad (1.48)$$

En définissant le coefficient de diffusion réduit  $\sigma = \sigma_s(1 - g)^1$ , l'équation suivante est finalement obtenue :

$$\frac{1}{c} \frac{\partial \mathbf{J}}{\partial t}(\mathbf{r}, t) + \frac{\nabla \varphi(\mathbf{r}, t)}{n_{\Omega}} = -(\kappa + \sigma) \mathbf{J}(\mathbf{r}, t) + S_{\mathbf{J}}(\mathbf{r}, t) \quad (1.49)$$

Afin d'obtenir le modèle de l'approximation diffuse, l'hypothèse de l'approximation de la diffusion doit être considérée. Cette hypothèse, qui consiste à supposer que le terme  $\frac{1}{c} \frac{\partial |\mathbf{J}|}{\partial t}(\mathbf{r}, t)$  est négligeable devant  $(\kappa + \sigma) |\mathbf{J}|(\mathbf{r}, t)$ , est justifiée lorsque la condition  $\kappa \ll \sigma$  est satisfaite (Arridge, 1999).

---

<sup>1</sup>Le coefficient de diffusion réduit  $\sigma$  peut-être vu comme un coefficient de diffusion isotrope équivalent. Son inverse définit le libre parcours moyen de transport et s'interprète comme la distance à partir de laquelle un photon a perdu toute information concernant sa direction initiale.

Sous l'hypothèse de l'approximation de la diffusion, on déduit de (1.49) que

$$\mathbf{J}(\mathbf{r}, t) = -D\nabla\varphi(\mathbf{r}, t) + \frac{S_{\mathbf{J}}(\mathbf{r}, t)}{\kappa + \sigma} \quad (1.50)$$

où le coefficient de diffusion macroscopique est défini tel que  $D = [n_{\Omega}(\kappa + \sigma)]^{-1}$ . L'AD est alors obtenue en insérant (1.50) dans (1.35) :

$$\frac{1}{c} \frac{\partial \varphi}{\partial t}(\mathbf{r}, t) - \nabla \cdot [D\nabla\varphi(\mathbf{r}, t)] + \kappa\varphi(\mathbf{r}, t) = S_{\varphi}(\mathbf{r}, t) - \nabla \cdot [n_{\Omega}DS_{\mathbf{J}}(\mathbf{r}, t)] \quad (1.51)$$

Dans ce manuscrit, le terme source de (1.51),  $S_{\varphi}(\mathbf{r}, t) - \nabla \cdot [n_{\Omega}DS_{\mathbf{J}}(\mathbf{r}, t)]$  est supposé nul. D'une part, aucune source volumique de rayonnement n'est supposée présente dans les milieux investigués. D'autre part, travaillant à température ambiante ( $\approx 300$  K) et dans le domaine spectral du visible / proche infrarouge, le phénomène d'émission thermique peut être négligé. En effet, d'après la loi de Wien, le flux maximal à cette température se situe dans l'infrarouge lointain ( $\approx 10 \mu\text{m}$ ).

### 1.4.3 Conditions de bord

Les conditions de bord de l'ÉTR étant directionnelles, l'approximation diffuse ne peut pas, à l'évidence satisfaire ces dernières. De par ce fait, les conditions de bord de l'approximation diffuse ont fait l'objet de nombreuses études dans la littérature. Il ressort de ces études trois conditions de bord couramment employées (Haskell *et al.*, 1994; Hielscher *et al.*, 1995; Schweiger *et al.*, 1995), que sont la condition de Dirichlet (ou condition nulle), la condition extrapolée et la condition de Robin. Les deux premières conditions de bord citées étant physiquement incorrectes et principalement utilisées en raison de leur simplicité mathématique (Haskell *et al.*, 1994; Hielscher *et al.*, 1995), ce texte présente ci-après, et considère par la suite, la condition de bord de Robin largement plébiscitée dans la littérature.



Soient  $n$  l'indice de réfraction du milieu d'étude  $\Omega$ ,  $n_a$  l'indice de réfraction du milieu ambiant,  $n \neq n_a$ , et les densités de flux radiatifs entrantes,  $\Phi_-$ , et sortantes,  $\Phi_+$ , définies par :

$$\Phi_-(\mathbf{r}, t) = - \int_{\{\mathbf{s} \cdot \mathbf{n} < 0\}} L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} \quad \mathbf{r} \in \partial\Omega \quad (1.52)$$

$$\Phi_+(\mathbf{r}, t) = \int_{\{\mathbf{s} \cdot \mathbf{n} > 0\}} L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} \quad \mathbf{r} \in \partial\Omega \quad (1.53)$$

La réflexion interne à l'interface entre les deux milieux contribue à la densité de flux radiatif entrante, qui s'écrit, en absence de sources lumineuses extérieures :

$$\Phi_-(\mathbf{r}, t) = \int_{\{\mathbf{s} \cdot \mathbf{n} > 0\}} R_1(\mathbf{r}, \mathbf{s}) L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} \quad \mathbf{r} \in \partial\Omega \quad (1.54)$$

Introduisant à présent un coefficient de réflexion effectif non directionnel  $R_{\text{eff}} = R_{\text{eff}}(\mathbf{r})$  vérifiant :

$$\int_{\{\mathbf{s} \cdot \mathbf{n} > 0\}} R_1(\mathbf{r}, \mathbf{s}) L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} = R_{\text{eff}} \int_{\{\mathbf{s} \cdot \mathbf{n} > 0\}} L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} \quad (1.55)$$

$$\Phi_-(\mathbf{r}, t) = R_{\text{eff}} \Phi_+(\mathbf{r}, t), \quad \mathbf{r} \in \partial\Omega \quad (1.56)$$

Insérant l'approximation P1, éq. (1.39), dans les éqs. (1.52) et (1.53), le résultat suivant peut-être démontré (Kaipio et Somersalo, 2005, p. 216) :

$$\Phi_{\pm}(\mathbf{r}, t) = \gamma_{n\Omega} \varphi(\mathbf{r}, t) \pm \frac{1}{2} \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n}, \quad \mathbf{r} \in \partial\Omega \quad (1.57)$$

où  $\gamma_2 = \frac{1}{\pi}$  et  $\gamma_3 = \frac{1}{4}$ . Finalement, insérant les éqs. (1.50) et (1.57) dans l'éq. (1.56), la condition de bord de type Robin suivante est obtenue :

$$\varphi(\mathbf{r}, t) + \frac{1}{2\gamma_{n\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D \nabla \varphi(\mathbf{r}, t) \cdot \mathbf{n} = 0, \quad \mathbf{r} \in \partial\Omega \quad (1.58)$$

Concernant la détermination du coefficient de réflexion effectif  $R_{\text{eff}}$ , (Groenhuis *et al.*, 1983) propose la relation expérimentale déterminée par moindres carrés :

$$R_{\text{eff}} \approx -1,4399n^{-2} + 0,7099n^{-1} + 0,6681 + 0,0636n \quad (1.59)$$

lorsque l'indice de réfraction du milieu ambiant est  $n_a = 1$ . (Haskell *et al.*, 1994) effectue le calcul théorique de  $R_{\text{eff}}$  à partir de la formule du facteur de réflexion de Fresnel, éq. (1.24). Ce dernier s'écrit :

$$R_{\text{eff}} = \frac{R_\varphi + R_{\mathbf{J}}}{2 - R_\varphi + R_{\mathbf{J}}} \quad (1.60)$$

avec

$$R_\varphi = \int_0^{\pi/2} 2 \sin(\theta) \cos(\theta) R_1(\theta) d\theta \quad (1.61)$$

$$R_{\mathbf{J}} = \int_0^{\pi/2} 3 \sin(\theta) \cos^2(\theta) R_1(\theta) d\theta \quad (1.62)$$

Enfin, (Keijzer *et al.*, 1988) considère une approximation constante par morceaux sur deux paliers de la fonction  $R_1(\theta)$  pour obtenir, de manière explicite, l'expression du coefficient  $A = \frac{1+R_{\text{eff}}}{1-R_{\text{eff}}}$  suivante :

$$A = \frac{\frac{2}{1-R_0} - 1 + |\cos(\theta_c)|^3}{1 - |\cos(\theta_c)|^2} \quad (1.63)$$

où  $R_0 = \left(\frac{n-n_a}{n+n_a}\right)^2$  et  $\theta_c = \arcsin(\frac{n_a}{n})$ . Par souci de simplicité, cette formulation sera considérée par la suite dans la condition de bord de l'AD, éq. (1.58).

À titre d'illustration, quelques valeurs du paramètre  $A$  pour les trois méthodes présentées ci-dessus et  $n_a = 1$  sont données dans le tableau 1.1. Il convient de noter que les valeurs de  $A$  selon (Haskell *et al.*, 1994) présentées dans ce tableau sont extraites de (Haskell *et al.*, 1994) et (Boffety, 2010). Enfin, il est intéressant de remarquer que les valeurs de  $A$  selon (Keijzer *et al.*, 1988) sont inférieures à celles fournies par (Haskell *et al.*, 1994) qui, comme le souligne (Boffety, 2010), sont remises en cause par (Chen et Bai, 1998) et seraient trop élevées.

TAB. 1.1 Quelques valeurs du paramètre  $A = \frac{1+R_{\text{eff}}}{1-R_{\text{eff}}}$  à l'interface entre deux milieux d'indices de réfraction différents selon (Groenhuis *et al.*, 1983), (Haskell *et al.*, 1994) et (Keijzer *et al.*, 1988) ( $n_a = 1$ ).

Références \ n	1,00	1,33	1,38	1,40	1,55
(Groenhuis <i>et al.</i> , 1983)	1,00	2,79	3,12	3,25	4,34
(Keijzer <i>et al.</i> , 1988)	1,00	2,35	2,63	2,74	3,71
(Haskell <i>et al.</i> , 1994)	1,00	2,51	2,82	2,94	3,99

#### 1.4.4 Quantité mesurable

En TO, la quantité mesurable est la densité de flux radiatif sortante qui, en prenant en compte la réflexion à l'interface, s'écrit :

$$\Phi_{\text{out}}(\mathbf{r}, t) = \int_{\{\mathbf{s} \cdot \mathbf{n} > 0\}} (1 - R_1(\mathbf{r}, \mathbf{s})) L(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} \quad (1.64)$$

$$= (1 - R_{\text{eff}}) \Phi_+(\mathbf{r}, t) \quad (1.65)$$

En appliquant les équations (1.57) puis (1.50), l'expression de l'équation (1.64) devient :

$$\Phi_{\text{out}}(\mathbf{r}, t) = (1 - R_{\text{eff}}) \gamma_{n_\Omega} \varphi(\mathbf{r}, t) - \frac{1}{2} (1 - R_{\text{eff}}) D \nabla \varphi(\mathbf{r}, t) \cdot \mathbf{n} \quad (1.66)$$

Mais d'après la condition de bord, éq. (1.58), l'expression suivante est valide :

$$2\gamma_{n_\Omega} (1 + R_{\text{eff}}) \varphi(\mathbf{r}, t) = - (1 - R_{\text{eff}}) D \nabla \varphi(\mathbf{r}, t) \cdot \mathbf{n} \quad (1.67)$$

Finalement, il apparaît que la quantité mesurable en TO est proportionnelle, sous l'hypothèse de l'approximation de la diffusion, à la densité de photons aux nœuds des détecteurs :

$$\Phi_{\text{out}}(\mathbf{r}, t) = 2\gamma_{n_\Omega} \varphi(\mathbf{r}, t) \quad (1.68)$$

### 1.4.5 Modélisation des sources

Dans le cadre de l'AD, deux modélisations des sources de lumière extérieures sont généralement considérées. Le premier modèle correspond au cas où une source diffuse est utilisée (DEL). Alors l'hypothèse de l'approximation P1 demeure valide près des sources et le terme source peut-être directement incorporé dans la condition de bord (1.58). Soit  $f$  la densité de flux de la source, définie telle que :

$$f(\mathbf{r}, t) = - \int_{\{\mathbf{s} \cdot \mathbf{n} < 0\}} \xi(\mathbf{r}, \mathbf{s}, t) \mathbf{s} \cdot \mathbf{n} d\mathbf{s} \quad \mathbf{r} \in \partial\Omega_s \quad (1.69)$$

En prenant en compte les réflexions au bord du domaine, l'équation (1.56) se réécrit :

$$\Phi_-(\mathbf{r}, t) = \begin{cases} R_{\text{eff}}\Phi_+(\mathbf{r}, t) + (1 - R_{\text{eff}})f(\mathbf{r}, t) & \mathbf{r} \in \partial\Omega_s \\ R_{\text{eff}}\Phi_+(\mathbf{r}, t) & \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.70)$$

d'où la condition de bord associée au modèle de la source diffuse suivante :

$$\varphi(\mathbf{r}, t) + \frac{1}{2\gamma_{n\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D \nabla \varphi(\mathbf{r}, t) \cdot \mathbf{n} = \begin{cases} \frac{f(\mathbf{r}, t)}{\gamma_{n\Omega}} & \mathbf{r} \in \partial\Omega_s \\ 0 & \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.71)$$

Le second modèle correspond à l'utilisation d'une source collimatée (diode laser). Dans ce cas, le terme source est généralement représenté par une source isotrope ponctuelle à l'intérieur du domaine  $\Omega$  :

$$q(\mathbf{r}, t) \delta(\mathbf{r} - \mathbf{r}_s), \quad \mathbf{r} \in \Omega \quad (1.72)$$

où  $\mathbf{r}_s$  est situé sous l'emplacement de la source à la distance  $(\kappa + \sigma)^{-1}$  (Patterson *et al.*, 1989; Haskell *et al.*, 1994; Hielscher *et al.*, 1995) ou  $\sigma^{-1}$  (Schweiger *et al.*, 1995; Ye *et al.*, 2001; Heino et Somersalo, 2002; Tarvainen, 2006; Azizi *et al.*, 2009). Le coefficient  $q$ , d'unité le  $\text{W m}^{-3}$ , s'interprète comme la quantité d'énergie radiative par unité de volume et par seconde provenant de la source extérieure. Il importe de spécifier que la méthode de calcul de ce coefficient à partir des caractéristiques de la source laser, de même que sa valeur considérée

dans le modèle direct, ne sont en général pas spécifiées dans la littérature. (Jiang, 2010, section 3.5) affirme qu'en principe  $q$  peut être mesuré, mais que cela est coûteux en temps et imprécis dans un contexte expérimental. (Jiang, 2010, section 3.5) propose par ailleurs une méthode de calibration pour déterminer ce paramètre expérimentalement. De plus, sa méthode de calibration inclut la détermination expérimentale des trois autres paramètres critiques, nécessaire pour l'obtention de reconstructions de haute qualité, que sont le coefficient  $A$  et les propriétés radiatives du fond (constantes). Enfin, d'autres implémentations des sources collimatées sont discutées dans la littérature, notamment les modélisations consistant à représenter le terme source sous la forme d'une ligne de densité de photons traversant le milieu et suivant une décroissance exponentielle (Eason *et al.*, 1978; Piron et L'Huillier, 2006).

Finalement, le système d'équations aux dérivées partielles requis par la suite de ces travaux pour une source diffuse est obtenu :

$$\frac{1}{c} \frac{\partial \varphi}{\partial t}(\mathbf{r}, t) - \nabla \cdot [D \nabla \varphi(\mathbf{r}, t)] + \kappa \varphi(\mathbf{r}, t) = 0 \quad \mathbf{r} \in \Omega \quad (1.73)$$

$$\varphi(\mathbf{r}, t) + \frac{1}{2\gamma_{n_\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D \nabla \varphi(\mathbf{r}, t) \cdot \mathbf{n} = \begin{cases} \frac{f(\mathbf{r}, t)}{\gamma_{n_\Omega}} & \mathbf{r} \in \partial\Omega_s \\ 0 & \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.74)$$

et pour une source collimatée :

$$\frac{1}{c} \frac{\partial \varphi}{\partial t}(\mathbf{r}, t) - \nabla \cdot [D \nabla \varphi(\mathbf{r}, t)] + \kappa \varphi(\mathbf{r}, t) = q(\mathbf{r}, t) \delta(\mathbf{r} - \mathbf{r}_s) \quad \mathbf{r} \in \Omega \quad (1.75)$$

$$\varphi(\mathbf{r}, t) + \frac{1}{2\gamma_{n_\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D \nabla \varphi(\mathbf{r}, t) \cdot \mathbf{n} = 0 \quad \mathbf{r} \in \partial\Omega \quad (1.76)$$

### 1.4.6 Passage au régime fréquentiel

Le passage du régime temporel au régime fréquentiel est réalisé à l'aide de la transformée de Fourier définie par :

$$\widehat{\varphi}(\mathbf{r}, \omega) = \int_{\mathbb{R}} \varphi(\mathbf{r}, t) \exp(-2i\pi\omega t) dt \quad (1.77)$$

$$\varphi(\mathbf{r}, t) = \int_{\mathbb{R}} \widehat{\varphi}(\mathbf{r}, \omega) \exp(2i\pi\omega t) d\omega \quad (1.78)$$

et des relations classiques :

$$\widehat{\frac{\partial \varphi}{\partial t}}(\mathbf{r}, \omega) = 2i\pi\omega \widehat{\varphi}(\mathbf{r}, \omega) \quad (1.79)$$

$$\widehat{\nabla \varphi}(\mathbf{r}, \omega) = \nabla \widehat{\varphi}(\mathbf{r}, \omega) \quad (1.80)$$

où  $\omega$  désigne ici la fréquence et  $\widehat{\varphi} : \Omega \times \mathbb{R} \rightarrow \mathbb{C}$ .

Le modèle de l'AD dans le régime fréquentiel s'écrit ainsi, pour une source diffuse :

$$-\nabla \cdot [D \nabla \widehat{\varphi}(\mathbf{r}, \omega)] + \left( \kappa + \frac{2i\pi\omega}{c} \right) \widehat{\varphi}(\mathbf{r}, \omega) = 0 \quad \mathbf{r} \in \Omega \quad (1.81)$$

$$\widehat{\varphi}(\mathbf{r}, \omega) + \frac{1}{2\gamma_{n_\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D \nabla \widehat{\varphi}(\mathbf{r}, \omega) \cdot \mathbf{n} = \begin{cases} \frac{\widehat{f}(\mathbf{r}, \omega)}{\gamma_{n_\Omega}} & \mathbf{r} \in \partial\Omega_s \\ 0 & \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.82)$$

et pour une source collimatée :

$$-\nabla \cdot [D \nabla \widehat{\varphi}(\mathbf{r}, \omega)] + \left( \kappa + \frac{2i\pi\omega}{c} \right) \widehat{\varphi}(\mathbf{r}, \omega) = \widehat{q}(\mathbf{r}, \omega) \delta(\mathbf{r} - \mathbf{r}_s) \quad \mathbf{r} \in \Omega \quad (1.83)$$

$$\widehat{\varphi}(\mathbf{r}, \omega) + \frac{1}{2\gamma_{n_\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D \nabla \widehat{\varphi}(\mathbf{r}, \omega) \cdot \mathbf{n} = 0 \quad \mathbf{r} \in \partial\Omega \quad (1.84)$$

Pour la suite de ce raisonnement, il est maintenant requis de supposer que les sources extérieures sont modulées en intensité à la fréquence  $\omega_0$ , par exemple de la forme :

$$\text{DC} + \text{AC} \times \cos(2\pi\omega_0 t) \quad (1.85)$$

où DC est l'intensité moyenne du signal, AC est l'amplitude de modulation et  $\omega_0$  est la fréquence de modulation. Puisque la transformée de Fourier de l'éq. (1.85) à la fréquence  $\omega_0$  vaut  $\frac{\text{AC}}{2}$ , il est possible de déduire l'équation décrivant l'évolution du champ de densité de photons à la fréquence  $\omega_0$ , pour une source diffuse :

$$-\nabla \cdot [D\nabla \hat{\varphi}(\mathbf{r}, \omega_0)] + \left( \kappa + \frac{2i\pi\omega_0}{c} \right) \hat{\varphi}(\mathbf{r}, \omega_0) = 0 \quad \mathbf{r} \in \Omega \quad (1.86)$$

$$\hat{\varphi}(\mathbf{r}, \omega_0) + \frac{1}{2\gamma_{n\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D\nabla \hat{\varphi}(\mathbf{r}, \omega_0) \cdot \mathbf{n} = \begin{cases} \frac{\text{AC}}{2\gamma_{n\Omega}} & \mathbf{r} \in \partial\Omega_s \\ 0 & \mathbf{r} \in \partial\Omega \setminus \partial\Omega_s \end{cases} \quad (1.87)$$

et pour une source collimatée :

$$-\nabla \cdot [D\nabla \hat{\varphi}(\mathbf{r}, \omega_0)] + \left( \kappa + \frac{2i\pi\omega_0}{c} \right) \hat{\varphi}(\mathbf{r}, \omega_0) = \frac{\text{AC}}{2} \delta(\mathbf{r} - \mathbf{r}_s) \quad \mathbf{r} \in \Omega \quad (1.88)$$

$$\hat{\varphi}(\mathbf{r}, \omega_0) + \frac{1}{2\gamma_{n\Omega}} \left( \frac{1 + R_{\text{eff}}}{1 - R_{\text{eff}}} \right) D\nabla \hat{\varphi}(\mathbf{r}, \omega_0) \cdot \mathbf{n} = 0 \quad \mathbf{r} \in \partial\Omega \quad (1.89)$$

Notons que le modèle de propagation du rayonnement, eqs. (1.86)-(1.87), sera utilisé dans les chapitres 3, 4 et 5 tandis que le modèle formé des équations (1.88)-(1.89) sera considéré dans les chapitres 2 et 6. La référence à la fréquence  $\omega_0$  sera généralement omise ainsi que la notation  $\hat{\cdot}$  représentative du domaine de Fourier. On précisera tout de même que  $\varphi$  est à valeurs complexes. Enfin, le terme  $\frac{\text{AC}}{2}$  sera remplacé par un simple terme source constant.

L'intérêt du régime fréquentiel réside dans la rapidité de résolution des équations (1.86)-(1.87) ou (1.88)-(1.89) par rapport au régime temporel, tout en permettant de collecter davantage d'information nécessaire à la résolution du problème inverse par rapport au régime permanent ( $\omega_0 = 0$ ). En effet, les systèmes d'équations aux dérivées partielles (1.86)-(1.87) et (1.88)-(1.89) s'apparentent à un système stationnaire à valeurs complexes, et ne sont donc pas beau-

coup plus compliqués et longs à résoudre que le système d'équations du régime permanent. D'autre part, le régime permanent ne permet d'obtenir qu'une mesure par détecteur, à savoir la diminution de l'intensité du signal après avoir traversé le milieu, tandis que le régime fréquentiel donne accès à la diminution de l'amplitude de modulation  $\left(\sqrt{\text{Re}(\varphi)^2 + \text{Im}(\varphi)^2}\right)$  ainsi qu'au déphasage ( $\arctan[\text{Im}(\varphi)/\text{Re}(\varphi)]$ ). Enfin, notons que la diminution de l'amplitude de modulation est déterminée conjointement par les phénomènes d'absorption et de diffusion tandis que le déphasage, qui reflète le temps de vol moyen des photons dans le milieu, est uniquement causé par la diffusion (Zhang, 2014). Une discussion sur la résolution des équations (1.88)-(1.89) est donnée dans le chapitre 2. Une validation numérique de la résolution des équations (1.86)-(1.87) et des ordres de convergence sont fournies en Annexe III.

#### 1.4.7 Domaine de validité

Les limitations de l'approximation de diffusion qu'il convient de ne pas perdre de vue sont récapitulées dans cette sous-section. D'abord, nous avons vu que le champ de luminance doit être supposé quasi-isotrope pour que l'hypothèse de l'approximation P1 soit valide. Cela implique qu'un photon doit subir suffisamment le phénomène de diffusion avant d'être absorbé et donc que  $\kappa \ll \sigma$ . Cette hypothèse ne peut être satisfaite près des sources collimatées, ce qui impose une distance suffisamment grande entre les sources et les détecteurs afin d'avoir des mesures de densité de photons correctes (une distance d'au moins 4 fois le libre parcours moyen de transport est recommandée par (O'Sullivan *et al.*, 2012) par exemple). Ensuite, nous avons vu que  $\kappa$  doit être négligeable devant  $\sigma$  afin que l'hypothèse de l'approximation de la diffusion, qui indique une faible variation temporelle du vecteur flux radiatif, soit justifiée (Arridge, 1999). Cela dit, cette hypothèse constitue une hypothèse de grande échelle de temps : dans le cas temporel, l'AD peut ne pas être valide aux temps courts (Elaloufi *et al.*, 2002; Francoeur *et al.*, 2005). Cela implique que même dans des situations où  $\kappa \ll \sigma$ , l'AD peut ne pas être précise pour des fréquences de modulation  $\omega$  trop grandes.

De nombreuses études ont été menées en comparant les résultats de l'AD avec ceux de l'ÉTR dans le but de déterminer le rapport  $\sigma/\kappa$  à partir duquel l'AD peut-être considérée valide.



Ce rapport est fixé à  $40\kappa \leq \sigma$  par (Hielscher et Alcouffe, 1996) tandis que (Bassani *et al.*, 1997; Nakai *et al.*, 1997) propose  $3\kappa \leq \sigma$ . En général, l'AD sera considérée valide dès lors que  $10\kappa \leq \sigma$  (O'Sullivan *et al.*, 2012). Cependant, il convient de noter que ce rapport ne peut être considéré sur l'ensemble des valeurs admissibles puisqu'il a été démontré que l'AD n'est pas valide dans des régions faiblement absorbantes et diffusantes (Firbank *et al.*, 1996; Hielscher *et al.*, 1998; Riley *et al.*, 2000; Dehghani *et al.*, 2000; Arridge *et al.*, 2000). Cette limitation pose problème notamment dans l'imagerie du cerveau en raison du liquide cérébro-spinal environnant le cerveau qui est très faiblement diffusant et absorbant, ou encore dans l'imagerie des articulations en raison de la présence du liquide synovial (Klose *et al.*, 2002). Il a de plus été observé que cette limitation pouvait entraîner d'importantes erreurs dans les reconstructions (Dehghani *et al.*, 2000). Afin de contourner les problèmes de validité de l'AD près des sources collimatées et dans les régions faiblement absorbantes et diffusantes, citons par d'exemple le développement de modèles de propagation de la lumière hybrides AD-ETR (Hayashi *et al.*, 2003; Tarvainen *et al.*, 2005a,b, 2006).

En conclusion, le domaine de validité de l'AD peut se résumer en supposant  $0 \ll \kappa \ll \sigma$ , une distance suffisamment grande entre les sources et les détecteurs (hypothèse de grande échelle spatiale) et des temps d'observation suffisamment grands.



## CHAPITRE 2

### ARTICLE 1 : REGULARIZATION OPPORTUNITIES FOR THE DIFFUSE OPTICAL TOMOGRAPHY PROBLEM

Fabien Dubot<sup>1,2</sup>, Yann Favennec<sup>2</sup>, Benoit Rousseau<sup>2</sup>, Daniel Rousse<sup>1</sup>

<sup>1</sup> Chaire de recherche industrielle t3e, École de technologie supérieure,

1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

<sup>2</sup> Laboratoire de thermocinétique de Nantes, UMR CNRS 6607,

Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France

Article soumis dans la revue « International Journal of Thermal Sciences » en décembre 2013.

Accepté en juin 2015.

#### Abstract

In optical tomography, the radiative properties of the medium under investigation are estimated from information contained in measurements provided by a set of light sources and sensors located on the frontier of the probed medium. Such a non-linear ill-posed inverse problem is usually solved through optimization with the help of gradient-type methods. Since it is well known that such inverse problems are ill-posed, regularization is required. This paper compares three distinct regularization strategies for two different optimization algorithms, namely the damped Gauss-Newton and BFGS algorithms, for the two-dimensional diffuse approximation to the radiative transfer equation in the frequency domain. More specifically, the mesh-based regularization is combined with the Tikhonov regularization in the damped Gauss-Newton algorithm. For the BFGS algorithm, the mesh-based regularization is combined with the Sobolev gradients method. Moreover, a space-dependent Sobolev gradients method is proposed for the first time. The performance of the proposed algorithms are compared by utilizing synthetic data. The deviation factor and correlation coefficient are used to quantitatively compare the final reconstructions. Also, three levels of noise are considered to characterize the behavior of the proposed methods against measurement noises. Numerical results indicate that the BFGS algorithm outperforms the damped Gauss-Newton in many aspects.

## 2.1 Introduction

In optical tomography (OT), the volumetric radiative properties of the medium under investigation are estimated from information contained in measurements provided by a set of light sources and sensors located on the frontier of the probed medium (Arridge, 1999; Charette *et al.*, 2008). The main applications of the OT concern the biomedical domain based on the fact that knowledge of the radiative properties provides information on the physiopathological condition of biological tissues. These applications include the imaging of finger joints (Hielscher *et al.*, 2004), breast imaging (Pogue *et al.*, 1997; Leff *et al.*, 2008), functional imaging of brain activity (Strangman *et al.*, 2002; Boas *et al.*, 2004) or small-animal imaging (Hielscher, 2005) for instance.

In order to looking deep into tissue, near-infrared light from about 650-950 nm is delivered through optical fibres because the absorption of light is relatively low in this spectral range (Klose *et al.*, 2002). However, the light is highly scattered at these wavelengths so that no direct methods can be employed to solve this problem. As a result, the reconstruction of the radiative property maps is usually performed through the solution of an inverse problem consisting in minimizing a cost function which depends on the discrepancy, in a least-square sense, between the measurements and associated predictions (Arridge, 1999). This method thus relies on a forward model that provides predictions assuming the distribution of radiative properties inside the medium is known.

The commonly accepted light propagation model in participating media is the radiative transfer equation (RTE) (Modest, 2013). This latter equation is integro-differential and thus heavy computations are needed to get accurate solutions. The computation may become highly CPU-time consuming when dealing with inverse problems for which solutions require large numbers of iterations. Alternatively, if the absorption coefficient of the medium is negligible compared to its scattering coefficient and if the medium does not contain void-like regions (i.e. regions in which the absorption and scattering are very low), then the diffuse approximation (DA) to the RTE provides a good approximation for describing light propagation in the medium (Arridge,

1999; Klose *et al.*, 2002). In this case one talks about the diffuse optical tomography (DOT) problem and the properties of interest are reduced scattering and absorption coefficients, denoted  $\sigma(\mathbf{r})$  and  $\kappa(\mathbf{r})$ , where  $\mathbf{r}$  is the space variable. The DA in the frequency domain is considered in this paper. Note that the nonuniqueness of the simultaneous reconstruction of absorption and reduced scattering coefficients in steady-state DOT has been demonstrated in (Arridge et Lionheart, 1998).

Many works have been undertaken over the past two decades to solve this non-linear ill-posed inverse problem, including the non-linear conjugate-gradient method (Arridge et Schweiger, 1998), Gauss-Newton (GN) based methods (Paulsen et Jiang, 1995; Schweiger *et al.*, 2005; Niu *et al.*, 2008; Dehghani *et al.*, 2009a) associated with a wavelet multi-scale method in (Zhu *et al.*, 1997a, 1998), shape-based reconstruction method (Zacharopoulos *et al.*, 2006; Arridge *et al.*, 2008) or, in a Bayesian framework, the approximation error method (Arridge *et al.*, 2006; Kaipio et Somersalo, 2005). While the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm was successfully used to solve the OT problem based on the RTE in (Klose et Hielscher, 2003; Ren *et al.*, 2006; Balima *et al.*, 2011a, 2013), it is rather surprising to note that its use was not considered, to the best knowledge of the authors, for the solution of the DOT problem (excepted in the very recent paper (Dubot *et al.*, 2015c) dedicated to the application of a wavelet multi-scale method associated with the BFGS method).

This paper is dedicated to a comparison between the BFGS algorithm and the damped GN algorithm when solving for the DOT problem. It is well known that such an inverse problem is ill-posed, thus regularization is mandatory. In this paper, three regularization tools are developed : i) mesh-based parameterization of the control variables (Chavent, 2010) ; ii) Tikhonov regularization ; and iii) Sobolev gradients method (Protas *et al.*, 2004). More specifically, the mesh-based regularization is combined with the Tikhonov regularization in the damped GN algorithm. For the BFGS algorithm, the mesh-based regularization is combined with the Sobolev gradients method. In addition, a comparison between two strategies is proposed, namely a gradient scaling and an adimensionalization of the control variables, to deal with the different orders of magnitude between the two distributions to be retrieved. In order to quantitatively

compared the proposed methods, deviation factor and correlation coefficient will be used in the numerical part to quantify the accuracy of the reconstructions. Moreover, in order to fully characterize the behavior of the proposed methods against measurement noises, three levels of noise will be considered.

Before presenting the work carried out in the paper, it should be noted that the efficiency of the damped GN algorithm to solve the DOT problem was shown in (Schweiger *et al.*, 2005). Also, it is worth mentioning that the mesh-based regularization has been studied in (Gu *et al.*, 2003) when solving for the DOT problem with the GN method associated with the Tikhonov regularization.

The novelties of the paper can be summarized as follow :

- The quasi-independence of the quasi-optimal Tikhonov parameter with the dimension of the control-space is shown and illustrated in the under-determined case (i.e. when the dimension of the control-space is greater than the number of measurements).
- A BFGS algorithm which combines together the mesh-based regularization and the Sobolev gradients method is designed to solve the DOT problem.
- A space-dependent Sobolev gradients method is developed and successfully applied for the solution of the DOT problem solved by the BFGS algorithm. This regularization tool is totally new, to the best knowledge of the authors, in the field of inversion.

The paper, which constitutes an extension of a previous work (Favennec *et al.*, 2013), is organized as follows : section 2 presents the forward model and its solution by the finite element method. This section also defines the cost function to be minimized. Section 3 brings mathematical tools on : i) strategies to deal with the different orders of magnitude between the two distributions to be retrieved ; ii) directional derivatives of first and second order of some functions ; and iii) finite element parameterization of the control variables. Section 4 presents in detail the optimization algorithms. Section 5 provides numerical results on a two-dimensional bounded domain to compare the different proposed algorithms. Finally, a detailed derivation of the cost function gradient with the adjoint-state method and through a continuous Lagrangian formulation is provided in Annex I.

As final comment about the current paper, it is worth mentioning that other types of penalization than the one of Tikhonov have been used when solving for the DOT problem with the GN method such as nonconvex and nonquadratic penalizations (Jagannath *et al.*, 2013; Shaw et Yalavarthy, 2014). The use of such penalization terms is out of the scope of the paper.

## 2.2 Problem statement

### 2.2.1 Forward model

The most commonly used forward model in OT is the DA to the RTE considered in the frequency domain. This is a simple equation governing the evolution of the photon density within the medium  $\Omega$ , say  $\phi$ . The DA model is written as (Arridge, 1999; Dehghani *et al.*, 2009a; Kaipio et Somersalo, 2005) :

$$\begin{aligned} -\nabla \cdot [D(\mathbf{r})\nabla\phi_l(\mathbf{r})] + \left[\kappa(\mathbf{r}) + \frac{2\pi i\omega}{c}\right] \phi_l(\mathbf{r}) &= q_l \delta(\mathbf{r} - \mathbf{r}_l) \quad \forall \mathbf{r} \in \Omega \\ \phi_l(\zeta) + \frac{A}{2\gamma_{n_\Omega}} D(\zeta) \nabla\phi_l(\zeta) \cdot \mathbf{n} &= 0 \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (2.1)$$

where  $\phi_l : \Omega \rightarrow \mathbb{C}$  is the state variable providing predictions for the  $l$ -th pointwise collimated light source,  $\mathbf{r} = (r_1, \dots, r_{n_\Omega})$  stands for the space variable,  $n_\Omega$  is the dimension of  $\Omega$ ,  $D(\mathbf{r}) = [n_\Omega (\kappa(\mathbf{r}) + \sigma(\mathbf{r}))]^{-1}$  is the macroscopic scattering coefficient (expressed in  $\text{cm}^{-1}$ ), where  $\kappa(\mathbf{r})$  and  $\sigma(\mathbf{r})$  are respectively the absorption and reduced scattering coefficients (both expressed in  $\text{cm}^{-1}$ ),  $q_l$  is the strength of the  $l$ -th source whose modulation frequency equals  $\omega$ ,  $\delta$  is the Dirac delta function,  $\mathbf{r}_l$  is located at the distance  $1/\sigma$  below the site of the  $l$ -th source which is located on the frontier of  $\Omega$ ,  $\partial\Omega$ ,  $i$  is the imaginary unit,  $c$  is the speed of light in the medium  $\Omega$ ,  $\mathbf{n}$  is the outer unit normal vector at  $\partial\Omega$ ,  $\gamma_{n_\Omega}$  is a dimension-dependent constant ( $\gamma_2 = 1/\pi$ ,  $\gamma_3 = 1/4$ ) (Kaipio et Somersalo, 2005) and  $A$  is a parameter that can be derived from the Fresnel laws if specular reflection is considered (Dehghani *et al.*, 2009a), or from experimental set-ups (Schweiger *et al.*, 1995).

This DA to the RTE is used when the specific intensity is assumed to be quasi-isotropic everywhere in the medium. A detailed description of the DA to the general RTE is given in (Arridge, 1999). It is well accepted that the DA model is a reasonably good approximation of the RTE as soon as the medium under consideration is highly scattering and satisfies  $0 \ll \kappa \ll \sigma$ .

The partial differential equation (2.1) is solved by the finite element method. To do so, let us define :

$$\hat{L}^2(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{C} \text{ such that } \int_{\Omega} |f(\mathbf{r})|^2 d\mathbf{r} < +\infty \right\} \quad (2.2)$$

$$\hat{H}^1(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{C} \text{ such that } f \in \hat{L}^2(\Omega) \text{ and } \frac{\partial f}{\partial r_j} \in \hat{L}^2(\Omega), \forall j = 1, \dots, n_{\Omega} \right\} \quad (2.3)$$

$$L^{\infty}(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \text{ such that } \exists C < +\infty : |f(\mathbf{r})| \leq C \text{ for almost every } \mathbf{r} \in \Omega\} \quad (2.4)$$

The variational formulation of equation (2.1) is (Schweiger *et al.*, 1995; Addam, 2009) : Find  $\varphi_l \in \hat{H}^1(\Omega)$ , such that

$$a(\varphi_l, w) = b(w) \text{ for all } w \in \hat{H}^1(\Omega) \quad (2.5)$$

where  $\bar{w} : \Omega \rightarrow \mathbb{C}$  is the complex conjugate of  $w$ , and the bilinear and linear forms are :

$$a(\varphi_l, w) = \int_{\Omega} D \nabla \varphi_l \cdot \nabla \bar{w} d\mathbf{r} + \int_{\Omega} \left[ \kappa + \frac{2\pi i \omega}{c} \right] \varphi_l \bar{w} d\mathbf{r} + \int_{\partial\Omega} \frac{2\gamma_{n_{\Omega}}}{A} \varphi_l \bar{w} d\zeta, \\ b(w) = \int_{\Omega} q_l \delta(\mathbf{r} - \mathbf{r}_l) \bar{w} d\mathbf{r} \quad (2.6)$$

Assuming  $\kappa, \sigma \in L^{\infty}(\Omega)$  and  $\kappa(\mathbf{r}), \sigma(\mathbf{r}) > 0$  for almost every  $\mathbf{r} \in \Omega$ , the existence and uniqueness of the solution  $\varphi_l \in \hat{H}^1(\Omega)$  to eq. (2.1) can be easily demonstrated through the complex version of the Lax-Milgram theorem (Brattka et Yoshikawa, 2006). Now, the discretization of the variational formulation, eq. (2.5), is addressed.

Let us consider a triangulation,  $\hat{\mathcal{M}}$ , of the computational domain  $\Omega$ . Let us note  $n_s$  the number of vertices in  $\hat{\mathcal{M}}$  and  $(\mathbf{r}^j)_{j=1}^{n_s}$  their corresponding coordinates. In this study, piecewise linear continuous functions are used to approximate the solution to eq. (2.5). These complex valued piecewise linear basis functions  $(\hat{\Phi}^i)_{i=1}^{n_s}$  are defined such that  $\hat{\Phi}^i$  is piecewise linear with respect



to the triangulation and  $\hat{\Phi}^i(\hat{\mathbf{r}}^j) = (1 + i)\delta_{ij}$  for all vertices  $\hat{\mathbf{r}}^j$  of  $\hat{\mathcal{M}}$ , where  $\delta_{ij}$  is the Kronecker delta ( $\delta_{ij} = 1$  if  $i = j$ ,  $\delta_{ij} = 0$  if  $i \neq j$ ). Let us denote  $\hat{\mathcal{V}}$  the finite dimensional subspace spanned by the functions  $(\hat{\Phi}^i)_{i=1}^{n_s}$ ,  $\hat{\mathcal{V}} \subset \hat{H}^1(\Omega)$ . Then,  $\varphi_l \in \hat{\mathcal{V}}$  is completely determined by its coordinate vector,  $\phi = (\phi_1, \dots, \phi_{n_s})^\top$ , relative to the basis  $(\hat{\Phi}^i)_{i=1}^{n_s}$  :

$$\varphi_l(\mathbf{r}) = \sum_{i=1}^{n_s} \varphi_l(\hat{\mathbf{r}}^i) \hat{\Phi}^i(\mathbf{r}) = \sum_{i=1}^{n_s} \phi_i \hat{\Phi}^i(\mathbf{r}) \quad (2.7)$$

In the same way, the real functions  $\kappa$ ,  $\sigma$  (and thus  $D = (n_\Omega(\kappa + \sigma))^{-1}$ ) are assumed to belong to the finite dimensional subspace spanned by the functions  $(\text{Re}(\hat{\Phi}^i))_{i=1}^{n_s}$ , where  $\text{Re}(\hat{\Phi}^i)$  denotes the real part of the complex basis function  $\hat{\Phi}^i$ . The radiative property maps are also completely determined by their coordinate vector,  $(\kappa(\hat{\mathbf{r}}^i))_{i=1}^{n_s}$  and  $(\sigma(\hat{\mathbf{r}}^i))_{i=1}^{n_s}$ , relative to the basis  $(\text{Re}(\hat{\Phi}^i))_{i=1}^{n_s}$  :

$$\kappa(\mathbf{r}) = \sum_{i=1}^{n_s} \kappa(\hat{\mathbf{r}}^i) \text{Re}(\hat{\Phi}^i)(\mathbf{r}) \quad ; \quad \sigma(\mathbf{r}) = \sum_{i=1}^{n_s} \sigma(\hat{\mathbf{r}}^i) \text{Re}(\hat{\Phi}^i)(\mathbf{r}) \quad (2.8)$$

Finally, the approximated variational formulation of eq. (2.5), which consists in finding  $\varphi_l \in \hat{\mathcal{V}}$  such that  $a(\varphi_l, w) = b(w)$  for all  $w \in \hat{\mathcal{V}}$ , leads to consider the approximated solution  $\varphi_l$  to eq. (2.1) through its coordinate vector  $\phi$  by solving the following linear algebraic system :

$$\mathbf{A}\phi = \mathbf{B} \quad (2.9)$$

where  $\mathbf{A} = (a_{ij})_{n_s \times n_s}$ ,  $\mathbf{B} = (b_i)_{n_s}$  and, for all  $i, j = 1, \dots, n_s$  :

$$a_{ij} = \int_{\Omega} D(\mathbf{r}) \nabla \hat{\Phi}^i \cdot \nabla \bar{\hat{\Phi}}^j \, d\mathbf{r} + \int_{\Omega} \left[ \kappa(\mathbf{r}) + \frac{2\pi i \omega}{c} \right] \hat{\Phi}^i \bar{\hat{\Phi}}^j \, d\mathbf{r} + \int_{\partial\Omega} \frac{2\gamma_{n_\Omega}}{A} \hat{\Phi}^i \bar{\hat{\Phi}}^j \, d\zeta \quad (2.10)$$

$$b_i = \int_{\Omega} q_l \delta(\mathbf{r} - \mathbf{r}_l) \bar{\hat{\Phi}}^i \, d\mathbf{r} \quad (2.11)$$

**Remark 1.** By definition of the Dirac delta function, the vector  $\mathbf{B}$  is equal to :

$$\mathbf{B} = \left( q_l \bar{\hat{\Phi}}^i(\mathbf{r} - \mathbf{r}_l) \right)_{i=1}^{n_s} \quad (2.12)$$

**Remark 2.** *In this paper, all computations are performed with the FreeFem++ environment (Hecht, 2012). The computation of the integrals involved in matrix  $\mathbf{A}$  and vector  $\mathbf{B}$  are carried out with the help of quadrature formulas which are exact for polynomials of degree 5. Moreover, among numerous linear system parallel solvers, the MULTifrontal Massively Parallel sparse direct Solver (MUMPS) (Amestoy et al., 2006) was chosen to solve the linear system, eq. (2.9), because it allows fast inversion, it is a direct solver and it is easily coupled with the FreeFem++ environment.*

### 2.2.2 Cost function

Let  $\partial\Omega_d = \{\hat{\mathbf{r}}_1, \dots, \hat{\mathbf{r}}_D\} \subset \{\hat{\mathbf{r}}^1, \dots, \hat{\mathbf{r}}^{n_s}\}$  be a set of  $D$  pointwise sensors on  $\partial\Omega$  and let us consider  $K$  collimated sources at different locations along  $\partial\Omega$ . For each source  $l \in \{1, \dots, K\}$ , which constitutes a test, let  $\check{\phi}_l : \Omega \rightarrow \mathbb{C}$  be equal to measurements of the amplitude and phase shift of the transmitted photon density. Predictions  $\phi_l = \phi_l^r + i\phi_l^i$  from the forward model (2.1) and measurements  $\check{\phi}_l = \check{\phi}_l^r + i\check{\phi}_l^i$  are integrated to the cost function to be minimized. Note that many authors consider a cost function  $\mathcal{J}(\phi) : \mathbb{C} \rightarrow \mathbb{R}$  but since some gradient-type algorithms shall be introduced later on, and since  $\mathcal{J}'(\phi)$  does not exist, i.e. this function is not holomorphic (the complex derivative of a real function does not exist, except for constant real functions), the differentiable cost function  $\mathcal{J}(\phi^r, \phi^i) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  that depends explicitly on both the real and imaginary parts of the density  $\phi$  is introduced. Hence, with no loss of generality, this yields the following cost function :

$$j(\gamma) := \mathcal{J}(\phi^r, \phi^i) := \frac{1}{2} \sum_{l=1}^K \sum_{j=1}^D \left( \frac{\phi_l^r(\hat{\mathbf{r}}_j) - \check{\phi}_l^r(\hat{\mathbf{r}}_j)}{\check{\phi}_l^r(\hat{\mathbf{r}}_j)} \right)^2 + \left( \frac{\phi_l^i(\hat{\mathbf{r}}_j) - \check{\phi}_l^i(\hat{\mathbf{r}}_j)}{\check{\phi}_l^i(\hat{\mathbf{r}}_j)} \right)^2 \quad (2.13)$$

The inverse problem can then be formulated as a constrained optimization problem of the form :

$$\gamma^* = \arg \min_{\gamma \in \mathcal{H} \times \mathcal{S}} j(\gamma) \text{ subject to } \phi_l \text{ solution of (2.1), } \forall l \in \{1, \dots, K\} \quad (2.14)$$

where  $\mathcal{H} = \mathcal{S} = L^\infty(\Omega)$ .

### 2.3 Mathematical settings

The two radiative properties to be retrieved from experiments are different in nature, and their order of magnitude also differ in the considered spectral range (near infra-red). As a consequence, the cost function gradient parts associated to both radiative properties also differ by roughly the same order of magnitude. In order to speed-up the iterative convergence to the local minimum, it is usual to follow (Balima *et al.*, 2011a; Klose, 2001) performing a scaling on the two parts of the computed cost function gradient :

$$\tilde{\nabla}_{\kappa} j(\gamma) = \chi_{\kappa} \nabla_{\kappa} j(\gamma), \quad \tilde{\nabla}_{\sigma} j(\gamma) = \chi_{\sigma} \nabla_{\sigma} j(\gamma) \quad (2.15)$$

with coefficients  $\chi_{\kappa}$  and  $\chi_{\sigma}$  computed before the first iteration with :

$$\chi_{\kappa} = 0.05 \frac{\|\kappa\|_{L^{\infty}(\Omega)}}{\|\nabla_{\kappa} j(\gamma)\|_{L^{\infty}(\Omega)}}, \quad \chi_{\sigma} = 0.05 \frac{\|\sigma\|_{L^{\infty}(\Omega)}}{\|\nabla_{\sigma} j(\gamma)\|_{L^{\infty}(\Omega)}}. \quad (2.16)$$

In this paper, the proposed strategy is different in the sense that the scaling is performed on the parameters themselves rather than on the cost function gradient. Choosing an a priori function for each radiative properties,  $\kappa_{ap}$  and  $\sigma_{ap}$  for the absorption and reduced scattering coefficients, respectively, one searches parameters that fluctuate about this prior. This adimensionalization leads to recover both  $\varkappa(\mathbf{r}) = \kappa(\mathbf{r})/\kappa_{ap}(\mathbf{r})$  and  $\varsigma(\mathbf{r}) = \sigma(\mathbf{r})/\sigma_{ap}(\mathbf{r})$  in respectively  $\mathcal{K}$  and  $\mathcal{S}$  and for which magnitude is of order one approximately for both coefficients. A comparison between results obtained with the use of this strategy which uses a priori functions  $\kappa_{ap}(\mathbf{r}) > 0$  and  $\sigma_{ap}(\mathbf{r}) > 0$  for all  $\mathbf{r} \in \Omega$ , and the gradient scaling strategy, is provided in the subsection 2.5.4. The notation  $\gamma = (\varkappa, \varsigma) \in \mathcal{K} \times \mathcal{S}$  is considered in the following.

Some complex inner products are also to be defined before calculations, integrating on the whole domain of interest  $\Omega$  or only on the boundaries  $\partial\Omega_d$  where the cost is integrated :

$$(f, g)_{\mathcal{X}} := \sum_{l=1}^K \sum_{j=1}^D f_l(\hat{\mathbf{r}}_j) \bar{g}_l(\hat{\mathbf{r}}_j) \quad ; \quad (f, g)_{\mathcal{Y}} := \sum_{l=1}^K \int_{\mathcal{D}} f_l \bar{g}_l \, d\mathbf{r} \quad (2.17)$$

where the subscript  $l$  refers to the source number and  $f$  and  $g$  are implicitly related to the complex functions  $f_l$  and  $g_l : \Omega \rightarrow \mathbb{C}$ ,  $l = 1, \dots, K$ .

The cost function, eq. (2.13), is rewritten in terms of the  $\hat{\mathcal{X}}$ -inner product as  $\mathcal{J}(\varphi^r, \varphi^i) = \frac{1}{2} (v(\varphi), v(\varphi))_{\hat{\mathcal{X}}}$  where  $\varphi = \varphi^r + i\varphi^i$  is implicitly used such that this cost function is actually written with real inner products in order to make the differentiation possible. Next, the relative error function is written as  $v(\varphi) = \frac{\varphi - \check{\varphi}}{\check{\varphi}}$  so that all different orders of magnitude are equivalently weighted within the cost function integration process. The cost function to be minimized, eq. (2.13), thus simply reads :

$$j(\gamma) = \frac{1}{2} \left( \frac{1}{|\check{\varphi}|^2} (\varphi - \check{\varphi}), \varphi - \check{\varphi} \right)_{\hat{\mathcal{X}}} \quad (2.18)$$

One also needs to define the directional derivatives  $j'(\gamma; \eta)$  and  $\varphi'(\gamma; \eta)$  at point  $\gamma \in \mathcal{K} \times \mathcal{S}$  and towards direction  $\eta \in \mathcal{K} \times \mathcal{S}$  as (Allaire et Craig, 2007) :

$$j'(\gamma; \eta) := \lim_{\varepsilon \rightarrow 0^+} \frac{j(\gamma + \varepsilon \eta) - j(\gamma)}{\varepsilon} \quad ; \quad \varphi'(\gamma; \eta) := \lim_{\varepsilon \rightarrow 0^+} \frac{\varphi(\gamma + \varepsilon \eta) - \varphi(\gamma)}{\varepsilon} \quad (2.19)$$

and extend this definition to the second order, according to (Lions et Faurre, 1982), with also  $\zeta \in \mathcal{K} \times \mathcal{S}$  such that :

$$j''(\gamma; \eta, \zeta) := \lim_{\varepsilon \rightarrow 0^+} \frac{j'(\gamma + \varepsilon \zeta; \eta) - j'(\gamma; \eta)}{\varepsilon} \quad (2.20)$$

The cost function being assumed to be twice differentiable at point  $\gamma$ , it can be shown that the directional derivative exists along direction  $\eta$  and the derivative to gradient and second derivative to Hessian relationships are then :

$$j'(\gamma; \eta) = (\nabla j(\gamma), \eta)_{L^2(\Omega)} \quad ; \quad j''(\gamma; \eta, \zeta) = (\nabla^2 j(\gamma) \eta, \zeta)_{L^2(\Omega)} \quad (2.21)$$

where the  $L^2$ -inner product is defined as  $(f, g)_{L^2(\Omega)} := \int_{\Omega} f g \, d\mathbf{r}$  for  $f$  and  $g : \Omega \rightarrow \mathbb{R}$ .

In order to use efficient optimization algorithms to solve eq. (2.14), the control-space must be approximate in order to be finite. Often, the finite element method is used so that one searches  $\kappa$  and  $\varsigma$  belonging to a finite dimensional subspace  $\mathcal{V}$ ,  $\mathcal{V} \subset \mathcal{K}$  and  $\mathcal{V} \subset \mathcal{S}$ . Let us consider a triangulation,  $\mathcal{M}$ , of the computational domain  $\Omega$ . Let us note  $n_p$  the number of vertices in  $\mathcal{M}$  and  $(\mathbf{r}^j)_{j=1}^{n_p}$  their corresponding coordinates. In this study, the prescriptions found in (Balima *et al.*, 2012) are followed in the sense that piecewise linear continuous functions are used to approximate  $\kappa$  and  $\varsigma$ . These piecewise linear basis functions  $(\Phi^i)_{i=1}^{n_p}$  are defined such that  $\Phi^i$  is piecewise linear with respect to the triangulation  $\mathcal{M}$  and  $\Phi^i(\mathbf{r}^j) = \delta_{ij}$  for all vertices  $\mathbf{r}^j$  of  $\mathcal{M}$ . Then,  $\mathcal{V}$  in the functional space spanned by these functions  $(\Phi^i)_{i=1}^{n_p}$  and the radiative properties  $\kappa$  and  $\sigma$  may be written as :

$$\begin{pmatrix} \kappa \\ \sigma \end{pmatrix}(\mathbf{r}) = \begin{pmatrix} \kappa_{ap} \\ \sigma_{ap} \end{pmatrix}(\mathbf{r}) \odot \sum_{\xi=1}^{n_p} \Phi^\xi(\mathbf{r}) \begin{pmatrix} \kappa \\ \varsigma \end{pmatrix}(\mathbf{r}^\xi) \quad (2.22)$$

where  $\odot$  is the element-wise vector product.

With the dimensionless parameterization (2.22), the discrete versions of eqs. (2.21) become :

$$j'(\gamma; \eta) = \eta^\top \nabla j(\gamma) \quad ; \quad j''(\gamma; \eta; \zeta) = \eta^\top \nabla^2 j(\gamma) \zeta \quad (2.23)$$

with  $\gamma \in \mathbb{R}_+^{2n_p}$ ,  $\eta$  and  $\zeta \in \mathbb{R}^{2n_p}$ ,  $\nabla j(\gamma) \in \mathbb{R}^{2n_p}$  and  $\nabla^2 j(\gamma) \in \mathbb{R}^{2n_p} \times \mathbb{R}^{2n_p}$ . In the following, we consider :

$$\nabla j(\gamma) = \begin{pmatrix} \nabla_{\kappa} j(\gamma) \\ \nabla_{\varsigma} j(\gamma) \end{pmatrix} \quad ; \quad \nabla^2 j(\gamma) = \begin{pmatrix} \nabla_{\kappa\kappa}^2 j(\gamma) & \nabla_{\kappa\varsigma}^2 j(\gamma) \\ \nabla_{\varsigma\kappa}^2 j(\gamma) & \nabla_{\varsigma\varsigma}^2 j(\gamma) \end{pmatrix} \quad (2.24)$$

**Remark 3.** *The consequences of the finite element parameterization of  $\kappa$  and  $\sigma$  in the solution of the DA equation, eq. (2.1), is pointed out. First, if the triangulations  $\mathcal{M}$  and  $\hat{\mathcal{M}}$  are the same, then the finite element basis  $(\Phi^i)_{i=1}^{n_p}$  and  $(\text{Re}(\hat{\Phi}^i))_{i=1}^{n_s}$  are also the same and no additional development is needed to solve the DA equation. Indeed, in this case, the values of  $\kappa$  and  $\sigma$  at coordinates  $(\hat{\mathbf{r}}^i)_{i=1}^{n_s}$  are known and the matrix  $\mathbf{A}$  can be built straightforwardly. However, a triangulation  $\mathcal{M}$  different from the triangulation  $\hat{\mathcal{M}}$ , with  $n_p < n_s$ , will be widely used in the rest of this paper, for regularization purposes. In this case, the values of  $\kappa$  and  $\sigma$  at co-*

ordinates  $(\hat{\mathbf{r}}^i)_{i=1}^{n_s}$  are obtained from those at coordinates  $(\mathbf{r}^i)_{i=1}^{n_p}$  through the fast finite element interpolation operator of the FreeFem++ environment (see chapter 6 of (Hecht et al., 2014)).

## 2.4 Optimization algorithms

### 2.4.1 Gauss-Newton algorithm

In the field of OT based on the DA model, the GN method and related methods (such as the Levenberg-Marquardt and others) are mostly used (Arridge, 1999; Schweiger *et al.*, 2005; Dehghani *et al.*, 2009a; Jagannath *et al.*, 2013). These optimization methods approximate the cost function Hessian and are therefore quasi-Newton-type optimization methods. More precisely, the cost function Hessian is approximated assuming that second-order state derivatives are negligible in comparison with the product of first-order state derivatives. Noting  $\tilde{\nabla}^2 j$  the approximate Hessian matrix, the GN matrix system is then written as :

$$\tilde{\nabla}^2 j(\gamma) \delta \gamma = -\nabla j(\gamma) \quad (2.25)$$

where  $\delta \gamma \in \mathcal{K} \times \mathcal{S}$  is the increment which has to be computed several times until convergence due to the non-linear behaviour of the forward model, eq. (2.1), with respect to properties  $\gamma = (\varkappa, \varsigma)$ . Starting from an initial estimate  $\gamma^0$ , the GN scheme solving for equation (2.14) is given by the iteration

$$\gamma^{m+1} = \gamma^m + \delta \gamma^m \quad (2.26)$$

where  $m$  denotes the iteration number.

Considering the cost function written in terms of the  $\hat{\mathcal{X}}$ -inner product, eq. (2.18), the gradient and approximate Hessian of the cost function are derived as follows :

$$\nabla j(\gamma) = \text{Re} \left( \mathbf{J}^\top \mathbf{R} \right) \quad ; \quad \tilde{\nabla}^2 j(\gamma) = \text{Re} \left( \mathbf{J}^\top \mathbf{J} \right) \quad (2.27)$$

where  $\mathbf{J} \in \mathbb{C}^{(K \times D) \times 2n_p}$  is the Jacobian matrix and  $\mathbf{R} \in \mathbb{C}^{(K \times D)}$  is the data model misfit whose expressions are obtained such that :

$$\mathbf{J} = \left( \frac{\varphi'_l(\boldsymbol{\kappa}; \Phi^i)(\hat{\mathbf{r}}_j)}{|\check{\varphi}_l(\hat{\mathbf{r}}_j)|}, \frac{\varphi'_l(\boldsymbol{\zeta}; \Phi^i)(\hat{\mathbf{r}}_j)}{|\check{\varphi}_l(\hat{\mathbf{r}}_j)|} \right)_{\substack{i=1, \dots, n_p \\ (l;j)=(1, \dots, K; 1, \dots, D)}} \quad (2.28)$$

$$\mathbf{R} = \left( \frac{\varphi_l(\hat{\mathbf{r}}_j) - \check{\varphi}_l(\hat{\mathbf{r}}_j)}{|\check{\varphi}_l(\hat{\mathbf{r}}_j)|} \right)_{(l;j)=(1, \dots, K; 1, \dots, D)} \quad (2.29)$$

The state derivatives  $\varphi'_l(\boldsymbol{\kappa}; \Phi^i)$  and  $\varphi'_l(\boldsymbol{\zeta}; \Phi^i)$  involved in equation (2.28) are computed differentiating the DA problem, eq. (2.1). For the direction  $\Phi^i$ ,  $i = 1, \dots, n_p$ , related to  $\boldsymbol{\kappa}$ , one obtains (Arridge, 1999) :

$$\begin{aligned} -\nabla \cdot [D \nabla \varphi'_l] + \left[ \boldsymbol{\kappa} + \frac{2\pi i \omega}{c} \right] \varphi'_l + \nabla \cdot [n_\Omega D^2 \boldsymbol{\kappa}_{ap} \Phi^i \nabla \varphi_l] + \boldsymbol{\kappa}_{ap} \Phi^i \varphi_l &= 0 \quad \forall \mathbf{r} \in \Omega \\ \varphi'_l + \frac{A}{2\gamma_{n_\Omega}} D \nabla \varphi'_l \cdot \mathbf{n} - \frac{A}{2\gamma_{n_\Omega}} n_\Omega D^2 \boldsymbol{\kappa}_{ap} \Phi^i \nabla \varphi_l \cdot \mathbf{n} &= 0 \quad \forall \boldsymbol{\zeta} \in \partial \Omega \end{aligned} \quad (2.30)$$

and for direction  $\Phi^i$ ,  $i = 1, \dots, n_p$ , related to  $\boldsymbol{\zeta}$  (Arridge, 1999) :

$$\begin{aligned} -\nabla \cdot [D \nabla \varphi'_l] + \left[ \boldsymbol{\kappa} + \frac{2\pi i \omega}{c} \right] \varphi'_l + \nabla \cdot [n_\Omega D^2 \boldsymbol{\sigma}_{ap} \Phi^i \nabla \varphi_l] &= 0 \quad \forall \mathbf{r} \in \Omega \\ \varphi'_l + \frac{A}{2\gamma_{n_\Omega}} D \nabla \varphi'_l \cdot \mathbf{n} - \frac{A}{2\gamma_{n_\Omega}} n_\Omega D^2 \boldsymbol{\sigma}_{ap} \Phi^i \nabla \varphi_l \cdot \mathbf{n} &= 0 \quad \forall \boldsymbol{\zeta} \in \partial \Omega \end{aligned} \quad (2.31)$$

Algorithm 2.1 gives some details about the construction of the vector  $\mathbf{R}$  and the matrix  $\mathbf{J}$ .

**Remark 4.** Equations (2.30) and (2.31) are akin to equation (2.1) with only different source terms in  $\Omega$  and on the boundary  $\partial \Omega$ , and thus, their solutions are not discussed. However, note that when solving for eqs. (2.30) and (2.31) to compute  $\varphi'_l(\boldsymbol{\kappa}; \Phi^i)$  and  $\varphi'_l(\boldsymbol{\zeta}; \Phi^i) \in \hat{\mathcal{V}}$ , respectively, projections of  $\boldsymbol{\kappa}$ ,  $\boldsymbol{\sigma}$ ,  $D$  and  $\Phi^i$ ,  $i = 1, \dots, n_p$ , from the finite element basis  $(\Phi^i)_{i=1}^{n_p}$  to the basis  $(\text{Re}(\hat{\Phi}^i))_{i=1}^{n_s}$  have to be performed if the triangulations  $\mathcal{M}$  and  $\hat{\mathcal{M}}$  are different.

**Remark 5.** Note that an adjoint method has been developed in (Arridge, 1999; Arridge et Schweiger, 1995) for the construction of the Jacobian matrix  $\mathbf{J}$ . This method is known to be computa-

Algorithme 2.1 : Construction of vector  $\mathbf{R}$  and matrix  $\mathbf{J}$ .

**Input :**  $\gamma = (\varkappa, \varsigma)$ ;

Compute  $\kappa = \kappa_{ap}\varkappa$ ,  $\sigma = \sigma_{ap}\varsigma$ ;

**for**  $l \leftarrow 1$  **to**  $K$  **do**

    Compute the state variable  $\varphi_l$  solving eq. (2.1);

    Extract the complex values of the photon density  $\varphi_l$  at the nodes of the sensors,  $\hat{\mathbf{r}}_j$ ,  $j = 1, \dots, D$ , and fill the corresponding part of the relative misfit to the source  $l$  in vector  $\mathbf{R}$ , eq. (2.29);

**for**  $i \leftarrow 1$  **to**  $n_p$  **do**

        Compute  $\varphi'_l(\varkappa; \Phi^i)$  and  $\varphi'_l(\varsigma; \Phi^i)$  solving eqs. (2.30) and (2.30), respectively;

        Extract the complex values of  $\varphi'_l(\varkappa; \Phi^i)$  and  $\varphi'_l(\varsigma; \Phi^i)$  at the nodes of the sensors,  $\hat{\mathbf{r}}_j$ ,  $j = 1, \dots, D$ , and fill the corresponding part to the source  $l$  and the direction  $\Phi^i$  in matrix  $\mathbf{J}$ , eq. (2.28);

**end**

**end**

**return**  $\mathbf{R}$ ,  $\mathbf{J}$ ;

*tionally efficient, especially if the number of measurements is lesser than the dimension of the control-space.*

It is well-known that the GN matrix system, eq. (2.25), is ill-posed and requires some regularization method. In order to obtain a stable solution while solving the optimization problem, eq. (2.14), by the GN algorithm, a penalty term involving the unknown radiative properties is added to the cost function, eq. (2.18). The Tikhonov regularization method, that has been employed extensively these last years together with the GN and related algorithms, is used in this study. This method consists in damping the matrix by adding the penalization  $\vartheta \mathcal{J}^+(\gamma)$  to the cost function  $\mathcal{J}(\varphi)$ , so that one searches the minimum of  $j^+(\gamma) = \mathcal{J}(\varphi) + \vartheta \mathcal{J}^+(\gamma)$ , where  $\vartheta$  is an appropriate regularization parameter. In the particular case where two property functions are to be estimated, it is convenient to penalize dimensionless properties. The additional term involved in the cost function thus reads :

$$\mathcal{J}^+(\gamma) = \frac{1}{2} \left\| \frac{\kappa - \kappa_{ap}}{\kappa_{ap}} \right\|_{L^2(\Omega)}^2 + \frac{1}{2} \left\| \frac{\sigma - \sigma_{ap}}{\sigma_{ap}} \right\|_{L^2(\Omega)}^2 \quad (2.32)$$



After basic calculations, this implies that the quantity  $\vartheta \frac{\kappa - \kappa_{ap}}{\kappa_{ap}}$  is added to the cost function gradient  $\nabla_{\kappa} j(\gamma)$ , and the quantity  $\vartheta \frac{\sigma - \sigma_{ap}}{\sigma_{ap}}$  is added to the cost function gradient  $\nabla_{\varsigma} j(\gamma)$ . Moreover, the quantity  $\vartheta$  is added to the diagonal of the matrix  $\tilde{\nabla}^2 j(\gamma)$ .

Numerous heuristic methods have been introduced to compute the quasi-optimal Tikhonov parameter  $\vartheta$ . Some methods use the assumed to be known standard deviation such as in the discrepancy principle (Morozov *et al.*, 1984; Goncharskii *et al.*, 1973). Others use less information about the noise properties, such as for instance the Generalized Cross Validation (GCV) (Golub *et al.*, 1979), or even no information about the noise level present in the system, such as the well known L-curve (Hansen et O’Leary, 1993; O’Leary, 2001). The latter method is used in this paper to determine the quasi-optimal Tikhonov parameter because of the simplicity of its implementation. Considering the under-determined case, i.e. when the dimension of the control-space is greater than the number of measurements, and a given level of noise, the quasi-independence of the quasi-optimal Tikhonov parameter with the control-space parameterization is illustrated in subsection 2.5.3.

Two common strategies may be adopted when updating the radiative properties with the GN method (Dennis et Schnabel, 1996). The first one is a trust region approach which consists in following the iterative scheme, eq. (2.26), computing the quasi-optimal Tikhonov parameter  $\vartheta$  at each iteration  $m$  before solving the GN matrix system, eq. (2.25). The second strategy is a damped GN approach which consists in computing the quasi-optimal Tikhonov parameter  $\vartheta$  only at the first iteration and performing a line-search for the descent direction  $\delta\gamma$  at each iteration. More precisely, the damped GN scheme solving for eq. (2.14) is given by the iteration

$$\gamma^{m+1} = \gamma^m + \alpha^* \delta\gamma^m \quad (2.33)$$

where the step-size  $\alpha^* \in \mathbb{R}_+$  is computed with the help of a line-search method which consists in solving the one-dimensional optimization problem :

$$\alpha^* \approx \arg \min_{\alpha > 0} j^+(\gamma^m + \alpha \delta\gamma^m) \quad (2.34)$$

These two strategies were compared in (Schweiger *et al.*, 2005). It is shown that the two approaches produce similar final reconstructions but the convergence is reached much more quickly with the damped GN. Hence, the latter approach was chosen in this study. The line-search method implemented in this paper is a quadratic interpolation search (algorithm 4.3 in (Antoniu et Lu, 2007)). However, note that the inexact line-search used in (Schweiger *et al.*, 2005) was also implemented, providing very similar results to those obtained with the quadratic interpolation search. The implementation of the damped GN method used in this paper is presented in algorithm 2.2. Finally, note that the linear algebraic system, eq. (2.25), is solved by the MUMPS solver (Amestoy *et al.*, 2006).

Algorithme 2.2 : The damped Gauss-Newton algorithm.

```

Input :  $m \leftarrow 1$ ,  $\gamma^0 = (\varkappa^0, \varsigma^0)$ ,  $\varkappa, \varsigma, N_M > 1$ ,  $\varepsilon_1, \vartheta$ ;

Compute  $j(\gamma^0)$ ;
Solve eq. (2.25) with  $\gamma = \gamma^0$  to compute  $\delta\gamma^0$ ;
 $\alpha^* \leftarrow \text{Line-Search}(\gamma^0, \delta\gamma^0, j(\gamma^0))$ ;
 $\gamma^1 \leftarrow \gamma^0 + \alpha^* \delta\gamma^0$ ;
Compute  $j(\gamma^1)$ ;
while  $m \leq N_M$  and  $\frac{|j(\gamma^m) - j(\gamma^{m-1})|}{j(\gamma^m)} > \varepsilon_1$  do
    Solve eq. (2.25) with  $\gamma = \gamma^m$  to compute  $\delta\gamma^m$ ;
     $\alpha^* \leftarrow \text{Line-Search}(\gamma^m, \delta\gamma^m, j(\gamma^m))$ ;
     $\gamma^{m+1} \leftarrow \gamma^m + \alpha^* \delta\gamma^m$ ;
    Compute  $j(\gamma^{m+1})$ ;
     $m \leftarrow m + 1$ ;
end
return  $\gamma^m$ ;

```

#### 2.4.2 Broyden-Fletcher-Goldfarb-Shanno algorithm

As a gradient-type method, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme solving for equation (2.14) is given by the iteration, eq. (2.33), starting from an initial estimate  $\gamma^0$ . At the iteration  $m$ , the descent direction  $\delta\gamma^m$  of the BFGS method is given by (Liu et Nocedal,

1989) :

$$\delta\gamma^m = -H_m^{-1}g_m \quad (2.35)$$

where  $g_m$  is the gradient of the cost function and  $H_m^{-1}$  is computed by the update formula :

$$H_{m+1}^{-1} = H_m^{-1} + \frac{s_m s_m^\top}{y_m^\top s_m} \left[ \frac{y_m^\top H_m^{-1} y_m}{y_m^\top s_m} + I_d \right] - \frac{1}{y_m^\top s_m} \left[ s_m y_m^\top H_m^{-1} + H_m^{-1} y_m s_m^\top \right] \quad (2.36)$$

where  $s_m = (\varkappa_{m+1} - \varkappa_m, \varsigma_{m+1} - \varsigma_m)$ ,  $y_m = g_{m+1} - g_m$  and  $H_0^{-1}$  is a positive-definite matrix (identity matrix in this study). According to (Nocedal, 1980), matrix  $H_m^{-1}$  shall not explicitly be computed in the algorithm for the sake of efficiency. Vectors  $s_m$  and  $y_m$  are stored during iterations and allow computation of the direction update, eq. (2.35).

A comparative study for the choice of line-search with the BFGS method has been realized and presented in (Dubot *et al.*, 2014). The golden-section search (algorithm 4.2 in (Antoniou et Lu, 2007)), the quadratic interpolation search (algorithm 4.3 in (Antoniou et Lu, 2007)) and an inexact line-search due to Fletcher (Fletcher, 1980, p. 126) have been considered. The conclusion is that the quadratic interpolation search is the most efficient among the three tested methods for the proposed application. The implementation of the BFGS method used in this paper is presented in algorithm 2.3.

The computation of the cost function gradient is now addressed. The adjoint-state method is used, which permits to access the cost function gradient at the low cost of only one additional adjoint problem. Let us define, for  $f_l$  and  $g_l : \Omega \rightarrow \mathbb{R}$ ,  $l = 1, \dots, K$ , the  $\mathcal{Y}$ -inner product as :

$$(f, g)_{\mathcal{Y}} := \sum_{l=1}^K \int_{\Omega} f_l g_l \, d\mathbf{r} \quad (2.37)$$

It can be shown that the directional derivatives of the cost function are equal to :

$$j'(\gamma; \eta) = (\text{Re}(\kappa_{ap} \varphi \bar{\varphi}^* - n_{\Omega} D^2 \kappa_{ap} \nabla \varphi \cdot \nabla \bar{\varphi}^*), \eta)_{\mathcal{Y}} \quad \forall \eta = (\eta_{\varkappa}, 0) \in \mathcal{K} \times \mathcal{S} \quad (2.38)$$

$$j'(\gamma; \eta) = (\text{Re}(-n_{\Omega} D^2 \sigma_{ap} \nabla \varphi \cdot \nabla \bar{\varphi}^*), \eta)_{\mathcal{Y}} \quad \forall \eta = (0, \eta_{\varsigma}) \in \mathcal{K} \times \mathcal{S} \quad (2.39)$$

Algorithme 2.3 : The BFGS algorithm.

**Input :**  $m \leftarrow 1$ ,  $\gamma^0 = (\varkappa^0, \varsigma^0)$ ,  $\varkappa$ ,  $\varsigma$ ,  $N_M > 1$ ,  $\varepsilon_1$ ;

Compute  $j(\gamma^0)$ ,  $\nabla j(\gamma^0)$ ;  
 $\delta\gamma^0 \leftarrow -\nabla j(\gamma^0)$ ;  
 $\alpha^* \leftarrow \text{Line-Search}(\gamma^0, \delta\gamma^0, j(\gamma^0))$ ;  
 $\gamma^1 \leftarrow \gamma^0 + \alpha^* \delta\gamma^0$ ;  
 Compute  $j(\gamma^1)$ ,  $\nabla j(\gamma^1)$ ;  
 $s(0) \leftarrow \gamma^1 - \gamma^0$ ;  $y(0) \leftarrow \nabla j(\gamma^1) - \nabla j(\gamma^0)$ ;  
**while**  $m \leq N_M$  and  $\frac{|j(\gamma^m) - j(\gamma^{m-1})|}{j(\gamma^m)} > \varepsilon_1$  **do**  
      $\delta\gamma^m \leftarrow \text{Direction-Update}(s, y, \nabla j(\gamma^m))$  following (Nocedal, 1980);  
      $\alpha^* \leftarrow \text{Line-Search}(\gamma^m, \delta\gamma^m, j(\gamma^m))$ ;  
      $\gamma^{m+1} \leftarrow \gamma^m + \alpha^* \delta\gamma^m$ ;  
     Compute  $j(\gamma^{m+1})$ ,  $\nabla j(\gamma^{m+1})$ ;  
      $s(m) \leftarrow \gamma^{m+1} - \gamma^m$ ;  $y(m) \leftarrow \nabla j(\gamma^{m+1}) - \nabla j(\gamma^m)$ ;  
      $m \leftarrow m + 1$ ;  
**end**  
**return**  $\gamma^m$ ;

in which  $\varphi_l^*$  is the solution to the adjoint problem :

$$\begin{aligned} -\nabla \cdot [D\nabla \varphi_l^*] + [\kappa - \frac{2\pi i \omega}{c}] \varphi_l^* &= 0 & \forall \mathbf{r} \in \Omega \\ \frac{2\gamma_{n\Omega}}{A} \varphi_l^* + D\nabla \varphi_l^* \cdot \mathbf{n} &= -\frac{1}{|\check{\varphi}_l|^2} (\varphi_l - \check{\varphi}_l) \mathbb{1}_{[\zeta \in \partial\Omega_d]}(\zeta) & \forall \zeta \in \partial\Omega \end{aligned} \quad (2.40)$$

where  $\mathbb{1}_{[\cdot]}$  denotes the indicator function, i.e.  $\mathbb{1}_{[\zeta \in \partial\Omega_d]}(\zeta) = 1$  if  $\zeta \in \partial\Omega_d = \{\hat{\mathbf{r}}_j\}_{j=1}^D$ ,  $\mathbb{1}_{[\zeta \in \partial\Omega_d]}(\zeta) = 0$ , otherwise.

Therefore, because the  $L^2$ -inner product is used to extract the cost function gradient in eq. (2.21), the expressions of the cost function gradients are such that :

$$\begin{aligned} \nabla_{\varkappa} j &= \sum_{l=1}^K \text{Re} \left( \kappa_{ap} \varphi_l \bar{\varphi}_l^* - n_{\Omega} D^2 \kappa_{ap} \nabla \varphi_l \cdot \nabla \bar{\varphi}_l^* \right) \\ \nabla_{\varsigma} j &= \sum_{l=1}^K \text{Re} \left( -n_{\Omega} D^2 \sigma_{ap} \nabla \varphi_l \cdot \nabla \bar{\varphi}_l^* \right) \end{aligned} \quad (2.41)$$

A proof of these results can either be obtained by use of Green functions (Arridge, 1999) or through a continuous Lagrangian formulation. This demonstration is carried out in Appendix I for the sake of completeness.

**Remark 6.** *Note that the adjoint problem, eq. (2.40), is very similar to the forward problem, eq. (2.1). Hence, its solution by the finite element method is explicitly not addressed. Moreover, note that  $\varphi_l$  and  $\varphi_l^*$  belong to the state-space, for all  $l = 1, \dots, K$ , while the cost function gradients, eq. (2.41), belongs to the control-space. As a result, let us consider the expansion of right-hand sides of eq. (2.41) in terms of the real and imaginary parts  $\text{Re}(\varphi_l)$ ,  $\text{Im}(\varphi_l)$ ,  $\text{Re}(\varphi_l^*)$ ,  $\text{Im}(\varphi_l^*)$ ,  $\text{Re}\left(\frac{\partial \varphi_l}{\partial r_1}\right)$ ,  $\dots$ ,  $\text{Im}\left(\frac{\partial \varphi_l^*}{\partial r_{n_\Omega}}\right)$ . In order to compute eq. (2.41) when the triangulations  $\mathcal{M}$  and  $\mathcal{M}$  are different, these real and imaginary parts are projected from the finite dimensional subspace spanned by the functions  $(\text{Re}(\hat{\Phi}^i))_{i=1}^{n_s}$  to the subspace  $\mathcal{V}$  with the help of the finite element interpolation operator.*

**Remark 7.** *Let us compare the computational cost per iteration for these two algorithms. Once the descent direction is obtained, the computational cost is identical for both algorithms. In terms of number of Partial Differential Equations (PDE) to solve, the cost function gradient requires the solution of  $(2 \times K)$  PDE for the BFGS algorithm while the Jacobian matrix requires the solution of  $(K + D)$  PDE if the adjoint method is used (Arridge, 1999). But above all, one observes that the descent direction of the BFGS algorithm is computed through the multiplication of the approximate inverse Hessian matrix with the cost function gradient, the latter matrix being computed in an explicit way. In contrast, for the GN algorithm, while the approximate Hessian matrix is computed in an explicit way from the Jacobian matrix, the descent direction is obtained through the solution of a large-scale linear algebraic system. The approximate Hessian matrix being full, the solution of this linear system can be CPU-time consuming. As a result, one can state that the BFGS algorithm outperforms the GN one from a computational cost point of view.*

## 2.5 Mixing regularization tools for Gauss-Newton and BFGS algorithms

### 2.5.1 Test presentation

The two-dimensional domain  $\Omega$  is bounded with a Cassini-type curve, centered at the origin, such that the diameter is approximately equal to 10 cm. 10 collimated pointwise sources and 32 pointwise sensors, equally spaced on the boundary  $\partial\Omega$ , are used to simulate the synthetic data (Fig. 2.1-f). The target properties to be reconstructed,  $\kappa_t$  and  $\sigma_t$ , are defined by (Figs. 2.1-g and 2.1-h) :

$$\begin{aligned}\kappa_t(\mathbf{r}) &= 0.1 \text{ cm}^{-1}, \quad \sigma_t(\mathbf{r}) = 30 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega_1, \quad \Omega_1 = \{(r_1 + 2)^2 + (r_2 + 0.75)^2 \leq 0.5^2\} \\ \kappa_t(\mathbf{r}) &= 0.06 \text{ cm}^{-1}, \quad \sigma_t(\mathbf{r}) = 10 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega_2, \quad \Omega_2 = \{(r_1 - 2)^2 + (r_2 - 0.75)^2 \leq 0.5^2\} \\ \kappa_t(\mathbf{r}) &= 0.08 \text{ cm}^{-1}, \quad \sigma_t(\mathbf{r}) = 20 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega \setminus (\Omega_1 \cup \Omega_2)\end{aligned}\tag{2.42}$$

The triangulation,  $\mathcal{M}$ , of the computational domain  $\Omega$  associated to the state variable is composed of 3 764 vertices (Fig. 2.1-b). Four triangulations,  $\mathcal{M}_j$ ,  $j = 1, \dots, 4$ , associated to the control variables  $\varkappa$  and  $\varsigma$  are considered to study the effect of the parameterization on the reconstructions obtained with the damped GN and the BFGS algorithms. The first triangulation considered,  $\mathcal{M}_1$ , is the same as the one associated to the state variable. Then, three coarser triangulations,  $\mathcal{M}_2$ ,  $\mathcal{M}_3$  and  $\mathcal{M}_4$ , are considered with 1 878, 978 and 170 vertices, respectively (Figs. 2.1-c, 2.1-d and 2.1-e).

The physical parameters involved in (2.1) are fixed to :  $\omega = 100 \times 10^6 \text{ s}^{-1}$ ,  $c = c_0/n_r$  where  $n_r = 1.4$  and  $c_0 = 3 \times 10^{10} \text{ cm.s}^{-1}$ ,  $\gamma_2 = \pi^{-1}$ ,  $A$  is derived from the Fresnel laws (see (Dehghani *et al.*, 2009a)) and  $q_l = 1 \text{ W.cm}^{-2}$  for all  $l = 1, \dots, 10$ . Note that the strength of the collimated sources has no impact on the final reconstructions.

Synthetic data are considered in this paper. These data,  $\check{\varphi}_l$ , representing the pseudo-experimental measurements, are built on a finer mesh,  $\hat{\mathcal{M}}_s$ , than that of generating the predictions  $\varphi_l$ ,  $l = 1, \dots, 10$ , in order to avoid the inverse crime (Kaipio et Somersalo, 2005; Colton et Kress, 1992)

(Fig. 2.1-a). Then, the finite element interpolation operator is used to obtain values of  $\check{\phi}_l$  at coordinates  $\hat{\mathbf{r}}_j, l = 1, \dots, 10, j = 1, \dots, 32$ . Finally, a multiplicative white gaussian noise is applied to the 320 synthetic data complex values  $(\check{\phi}_l(\hat{\mathbf{r}}_j))_{l=1, \dots, 10}^{j=1, \dots, 32}$  to simulate the noise inherent to experimental devices. To do so, being given the sequence  $(m_l^j)_{l=1, \dots, 10}^{j=1, \dots, 32}$  where  $m_l^j \sim \mathcal{N}(0, 1)$ , the noisy data are computed through the formula (Klose, 2001) :

$$\check{\phi}_l^{\text{noisy}}(\hat{\mathbf{r}}_j) = \check{\phi}_l(\hat{\mathbf{r}}_j) \left( 1 + 10^{-\frac{\text{SNR}}{10}} \times m_l^j \right) \quad (2.43)$$

where SNR denotes the signal-to-noise ratio, expressed in dB. In order to characterize the behaviour of the proposed inverse methods against noises, SNR values of 30, 20 and 10 dB are considered in the following.

Note that dimensions of the control space associated with meshes  $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$  and  $\mathcal{M}_4$  are 7 528, 3 756, 1 956 and 340, respectively. As a result, the synthetic data involving 640 real measurements, the inverse problems associated to meshes  $\mathcal{M}_1, \mathcal{M}_2$  and  $\mathcal{M}_3$  are under-determined while the one associated to the mesh  $\mathcal{M}_4$  is over-determined.

Finally, the inputs of the damped GN and BFGS methods, algorithms 2.2 and 2.3, respectively, are the following : (i) the dimensionless properties are initialized to unity, i.e.  $\gamma^0 = 1$ , while the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  are fixed to the background values, i.e.  $\kappa_{ap}(\mathbf{r}) = 0.08$  and  $\sigma_{ap}(\mathbf{r}) = 20$  for all  $\mathbf{r} \in \Omega$  ; (ii) the maximal number of iterations is set to 100, i.e.  $N_M = 100$  ; and (iii)  $\varepsilon_1 = 10^{-2}$ . Note that inputs  $N_M$  and  $\varepsilon_1$  have been chosen so as to provide reconstructions as accurate as possible for both algorithms. In particular, if a higher maximum iteration number and a lower value of  $\varepsilon_1$  are given, the cost function could continue to decrease but the quality of reconstructions will deteriorate.

## 2.5.2 Definition of reconstruction errors

In order to gauge the accuracy of the reconstructions, two errors are introduced. Let  $\lambda^j$  and  $\lambda_t^j$  be either the reconstructed and target absorption or reduced scattering coefficients at the node  $j$ , respectively, and  $n_s$  be the dimension of the state-space. Because several parameterizations

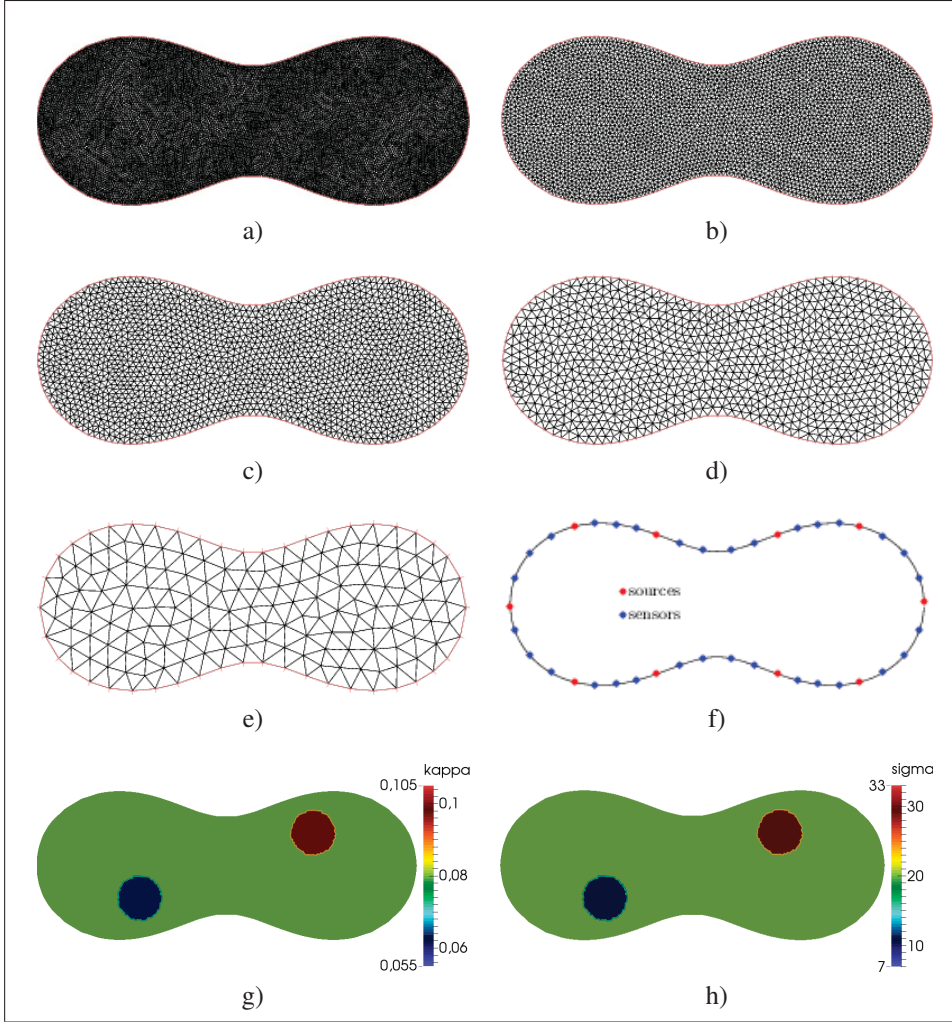


FIG. 2.1 Synthetic data mesh : (a)  $\hat{\mathcal{M}}_s$ , 13 143 nodes ; (b)  $\hat{\mathcal{M}} = \mathcal{M}_1$ , 3 764 nodes ; (c)  $\mathcal{M}_2$ , 1 878 nodes ; (d)  $\mathcal{M}_3$ , 978 nodes ; (e)  $\mathcal{M}_4$ , 170 nodes ; (f) locations of sources and sensors ; (g) maps of  $\kappa_t$  ; (h) maps of  $\sigma_t$ .

of the control-space will be used in the following and in order not to introduce bias, it should be noted that projections of the radiative property maps from the finite element basis  $(\Phi^i)_{i=1}^{n_p}$  to the finite element basis  $(\text{Re}(\hat{\Phi}^i))_{i=1}^{n_s}$  will be carried out after the convergence of the algorithms and before the computation of errors.



The first error is the deviation factor,  $e_1^\lambda$ , which measures the deviation of the reconstructed image from the target image. This error is defined as :

$$e_1^\lambda = 100 \times \frac{\sqrt{\sum_{j=1}^{n_s} (\lambda^j - \lambda_t^j)^2}}{\sqrt{\sum_{j=1}^{n_s} (\lambda_t^j)^2}} \quad (2.44)$$

The second error is the correlation coefficient (Klose et Hielscher, 2003; Flannery *et al.*, 1992),  $e_2^\lambda$ , which measures the linear correlation between the target and the reconstructed image. This error is defined as :

$$e_2^\lambda = \frac{\sum_{j=1}^{n_s} (\lambda^j - \bar{\lambda}) (\lambda_t^j - \bar{\lambda}_t)}{(n_s - 1) \sigma_\lambda \sigma_{\lambda_t}} \quad (2.45)$$

where  $\bar{\lambda}$  and  $\bar{\lambda}_t$  are the mean values of vectors  $\lambda$  and  $\lambda_t$  and  $\sigma_\lambda$ ,  $\sigma_{\lambda_t}$  are standard deviations of target and reconstructed images, respectively. Standard deviations are given by :

$$\sigma_\lambda = \sqrt{\frac{\sum_{j=1}^{n_s} (\lambda^j - \bar{\lambda})^2}{n_s - 1}} \text{ and } \sigma_{\lambda_t} = \sqrt{\frac{\sum_{j=1}^{n_s} (\lambda_t^j - \bar{\lambda}_t)^2}{n_s - 1}} \quad (2.46)$$

**Remark 8.** A small value of  $e_1^\lambda$  indicates a reconstructed image with high accuracy. A large value of  $e_2^\lambda$  shows a high detectability in the reconstructed image and indicates a reconstructed image with high accuracy.

### 2.5.3 Gauss-Newton algorithm : mixing Tikhonov penalization and mesh-based parameterization

As indicated in subsection 2.4.1, the regularization parameter  $\vartheta$  has to be determined at the beginning of the optimization process for the damped GN algorithm. In this paper, the L-curve method (Hansen et O’Leary, 1993; O’Leary, 2001) is used to determine an appropriate regularization parameter  $\vartheta$ . This method consists in plotting the solution norm,  $\mathcal{J}^+$ , versus the residual norm,  $\mathcal{J}$ , solving the GN matrix system for different values of the regularization parameter, and choosing  $\vartheta$  from the point on this curve that has maximum curvature, i.e. the

corner. The Reginska's method (Reginska, 1996) is used to estimate this corner and thus the quasi-optimal Tikhonov parameter. This method consists in locating the corner as the minimum point of the  $(\pi/4)$ -rotated L-curve.

The L-curves and  $(\pi/4)$ -rotated corresponding curves for levels of noise of 30 dB and 10 dB, and for parameterizations associated to meshes  $\mathcal{M}_1$ ,  $\mathcal{M}_2$  and  $\mathcal{M}_3$  are depicted in Fig. 2.2. It is observed from Figures 2.2-a and 2.2-b that the quasi-optimal Tikhonov parameter for a 30 dB noise equals 0.9 for all the three parameterizations. From Figures 2.2-c and 2.2-d, it is seen that the appropriate Tikhonov parameter for a 10 dB noise equals 1 for all the three parameterizations. Note that the case of a 20 dB noise is not depicted because its L-curves are very similar to those for the 30 dB case. In the over-determined case, i.e. when the radiative properties are reconstructed on the mesh  $\mathcal{M}_4$  (170 nodes), the proposed method has given an appropriate Tikhonov parameter equal to 1.5 for the three considered levels of noise. This illustrates that, for a given level of noise, the quasi-optimal Tikhonov parameter is quasi-independent of the control-space parameterization in the case of under-determined inverse problems. A more detailed study on this subject was carried on in Dubot *et al.* (2015a).

Now that the Tikhonov parameter has been determined for all parameterizations and all levels of noise, the damped GN method, algorithm 2.2, is applied. Table 2.1 gives errors in matrix form  $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$ , for the 12 considered cases. Results show that, for a given mesh among meshes  $\mathcal{M}_1$ ,  $\mathcal{M}_2$  and  $\mathcal{M}_3$ , errors are very similar whatever the level of noise. For the mesh  $\mathcal{M}_4$  and the 10 dB case, errors  $e_2$  are smaller and errors  $e_1$  are larger than those of the 20 dB and 30 dB cases. For the mesh  $\mathcal{M}_4$  and the 20 dB and 30 dB cases, errors are very close. Overall, it is observed that, in the under-determined case, the damped GN algorithm is very few sensitive to the level of noise. In the over-determined case, the damped GN algorithm is a little bit sensitive to the level of noise : final reconstructions are better for a 30 dB noise than for a 10 dB noise.

For a given level of noise, Table 2.1 shows that a reduction of the control-space dimension leads to improved results. The best reconstructions are thus obtained in the over-determined case. In order to assess these results and to illustrate how an appropriate parameterization of the

control-space can enhance the quality of reconstructions, the reconstructed images are depicted in Figures 2.3 and 2.4 for the 30 dB and 10 dB cases, respectively. As expected, the coarser the mesh, the better the reconstructions. Because of similar final results for the 20 dB and 30 dB cases, the reconstructions obtained for a 20 dB level of noise are not given.

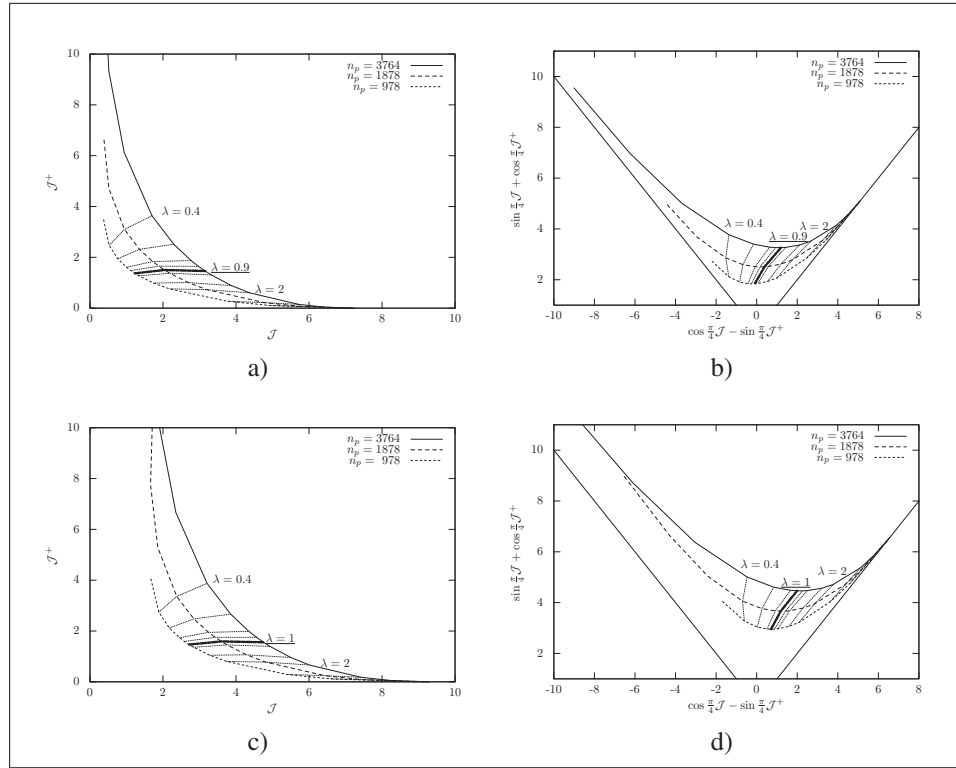


FIG. 2.2 L-curves (left) and  $(\pi/4)$ -rotated corresponding curves (right) for a 30 dB noise (top) and 10 dB noise (bottom). All curves are built for the three meshes  $\mathcal{M}_1$ ,  $\mathcal{M}_2$  and  $\mathcal{M}_3$  where  $n_p = 3764$ , 1878 and 978, respectively. Presented Tikhonov parameters are equal to 0.4, 0.6, 0.8, 0.9, 1, 1.1, 1.5, 2, 5, 10 and 100.

**Remark 9.** *It should be noted that the result on the quasi-independence of the quasi-optimal Tikhonov parameter with the control-space parameterization for under-determined inverse problems may seem unnecessary in the present study because best results for the damped GN algorithm are obtained on the coarsest mesh. However, because the construction of the Jacobian matrix can be very time consuming, this result is important for less ill-posed inverse problems for which a reduction of the dimension of the control-space is not of crucial importance. In this*

TAB. 2.1 Gauss-Newton algorithm. Obtained reconstruction errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for 4 meshes and 3 levels of noise.

Mesh \ Noise	10 dB	20 dB	30 dB
$\mathcal{M}_1$	$\begin{pmatrix} 5.43 & 11.08 \\ 0.61 & 0.56 \end{pmatrix}$	$\begin{pmatrix} 5.43 & 11.08 \\ 0.61 & 0.57 \end{pmatrix}$	$\begin{pmatrix} 5.43 & 11.08 \\ 0.61 & 0.57 \end{pmatrix}$
$\mathcal{M}_2$	$\begin{pmatrix} 5.10 & 10.39 \\ 0.64 & 0.63 \end{pmatrix}$	$\begin{pmatrix} 5.09 & 10.39 \\ 0.64 & 0.64 \end{pmatrix}$	$\begin{pmatrix} 5.09 & 10.39 \\ 0.64 & 0.64 \end{pmatrix}$
$\mathcal{M}_3$	$\begin{pmatrix} 4.80 & 9.62 \\ 0.67 & 0.68 \end{pmatrix}$	$\begin{pmatrix} 4.78 & 9.58 \\ 0.67 & 0.71 \end{pmatrix}$	$\begin{pmatrix} 4.78 & 9.58 \\ 0.67 & 0.71 \end{pmatrix}$
$\mathcal{M}_4$	$\begin{pmatrix} 4.29 & 7.84 \\ 0.73 & 0.78 \end{pmatrix}$	$\begin{pmatrix} 4.19 & 7.58 \\ 0.75 & 0.81 \end{pmatrix}$	$\begin{pmatrix} 4.19 & 7.58 \\ 0.75 & 0.82 \end{pmatrix}$

case, the quasi-optimal Tikhonov parameter should be identified based on a coarse mesh for the control-space, even though a finer mesh is to be used afterwards for better reconstructions.

#### 2.5.4 BFGS algorithm : comparison between adimensionalization and gradient scaling strategies

As explained in the section 2.3, two strategies may be adopted to deal with the different orders of magnitude between the radiative properties  $\kappa$  and  $\sigma$  when using the BFGS method, algorithm 2.3, namely a scaling of the cost function gradients or an adimensionalization of the radiative properties. These two strategies are compared, applying the BFGS algorithm for the three levels of noise considered and a parameterization of the radiative properties based on the mesh  $\mathcal{M}_1$ . The reconstruction errors for these six tests are reported in Table 2.2. It is seen from this table that, among the twelve computed errors for each strategy, ten of them are better for the adimensionalization of the radiative properties and one of them gives the same error. Figures 2.5 and 2.6 show the final reconstructions obtained for a 10 dB and 30 dB level of noise, respectively. Overall, it is clear from these figures that the adimensionalization strategy provides the best results.

Finally, let us remark the high fluctuations in the reconstructions near the boundary of the domain. This comes from the fact that the noise propagates from the data to the reconstructions

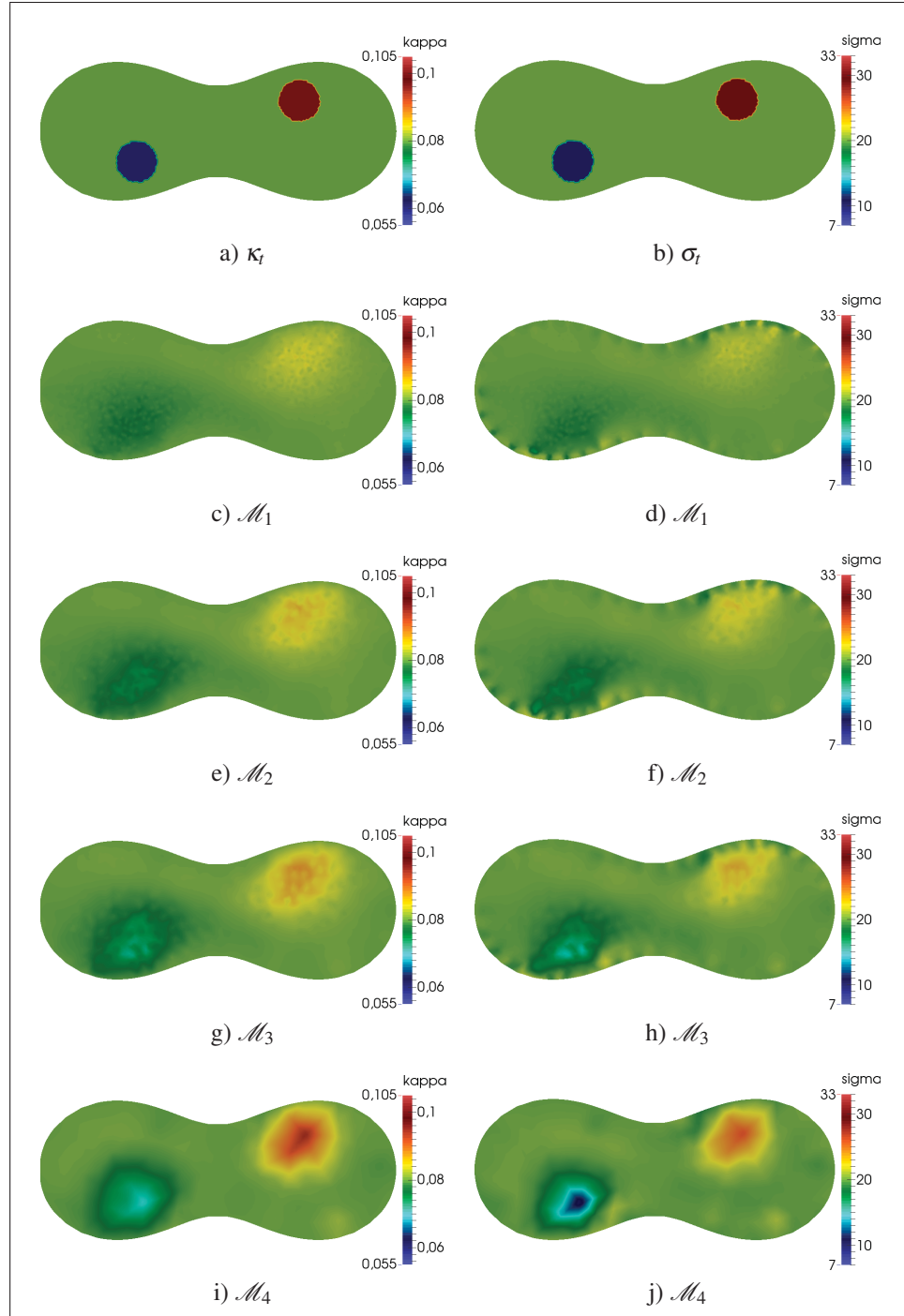


FIG. 2.3 Gauss-Newton algorithm ; 30 dB noise.  
Reconstructions for 4 control-space parameterizations. Left :  $\kappa$ ,  
right :  $\sigma$ . First raw : targets, next raws : reconstructions.

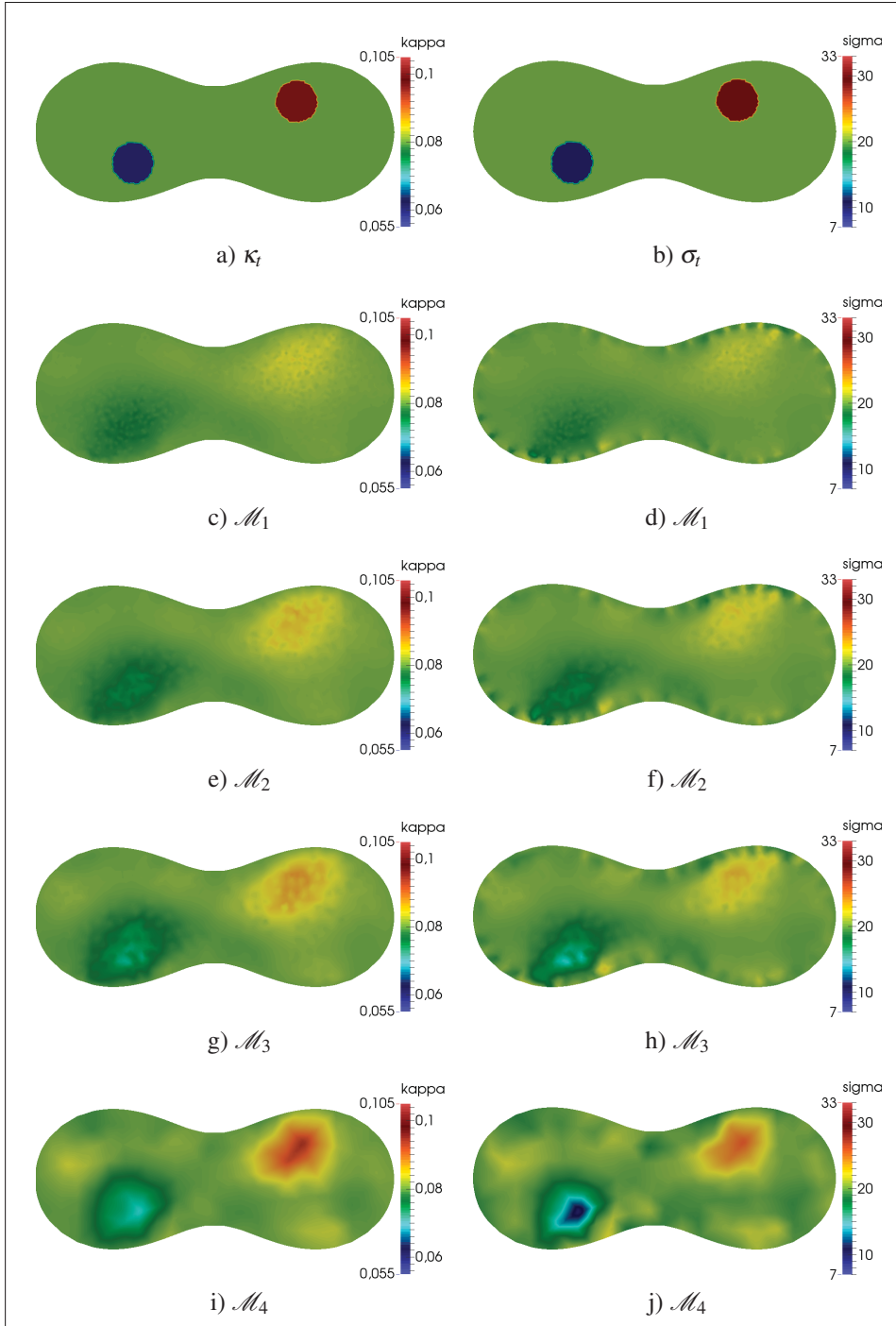


FIG. 2.4 Gauss-Newton algorithm ; 10 dB noise.  
Reconstructions for 4 control-space parameterizations. Left :  $\kappa$ ,  
right :  $\sigma$ . First raw : targets, next raws : reconstructions.

through the following path : noisy adjoint states are obtained because of the boundary source term  $-\frac{(\varphi_l - \check{\varphi}_l)}{|\check{\varphi}_l|^2} \mathbb{1}_{[\zeta \in \partial\Omega_d]}(\zeta)$  involved in the adjoint model (2.40) which contains the data that is indeed noisy. Then, noisy cost function gradients, eq. (2.41), are obtained because they depend on noisy adjoint states. Eventually the re-actualization process of the properties contains (at least implicitly) the noise present within the data.

TAB. 2.2 BFGS algorithm ; mesh  $\mathcal{M}_1$ . Obtained reconstruction errors  $\begin{pmatrix} e_1^k & e_1^\sigma \\ e_2^k & e_2^\sigma \end{pmatrix}$  for both the gradient scaling approach and the adimensionalisation approach. 3 levels of noise are considered.

Method \ Noise	10 dB	20 dB	30 dB
<b>Scaling</b>	$\begin{pmatrix} 5.32 & 9.08 \\ 0.77 & 0.68 \end{pmatrix}$	$\begin{pmatrix} 4.11 & 7.52 \\ 0.82 & 0.80 \end{pmatrix}$	$\begin{pmatrix} 4.08 & 7.49 \\ 0.82 & 0.81 \end{pmatrix}$
<b>Adim.</b>	$\begin{pmatrix} 4.22 & 8.95 \\ 0.73 & 0.69 \end{pmatrix}$	$\begin{pmatrix} 3.57 & 6.76 \\ 0.82 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 3.56 & 6.63 \\ 0.83 & 0.84 \end{pmatrix}$

### 2.5.5 BFGS algorithm : mesh-based parameterization

The effects of the mesh-based parameterization of the radiative properties on the final reconstructions when using the BFGS algorithm are investigated in this subsection. Table 2.3 gives the reconstruction errors for all the twelve tests investigated. In contrast to the damped GN algorithm, it is observed that for a given mesh, the BFGS algorithm is a little bit sensitive to the SNR : the  $e_1$  errors decrease when the SNR increases while the  $e_2$  errors increase when the SNR increases. Moreover, it is seen that for a given SNR, the  $e_1$  errors decrease when the dimension of the control-space decreases. However, when the dimension of the control-space is too small, i.e. when the radiative property meshes is too coarse, the radiative properties cannot be reconstructed with high accuracy and  $e_1$  errors increase. In the same way, the  $e_2$  errors increase when the dimension of the control-space decrease. However, when the dimension of the control-space is too small,  $e_2$  errors decreases. To illustrate that the reduction of the control-space dimension, until a certain limit, improves the final reconstructions obtained with the BFGS algorithm, Figures 2.7 and 2.8 present these reconstructions for a SNR equal to 30

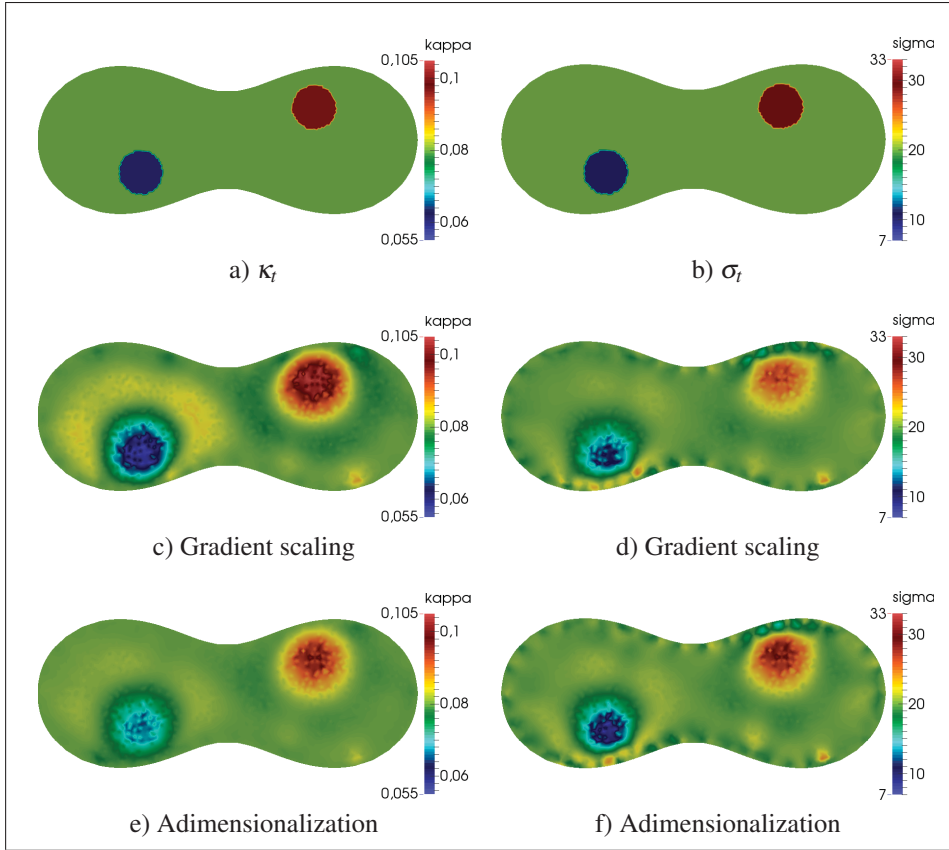


FIG. 2.5 BFGS algorithm ; mesh  $\mathcal{M}_1$  ; 30 dB noise.  
Reconstructions for both the gradient scaling approach and the  
adimensionalisation approach. Left :  $\kappa$ , right :  $\sigma$ . First row :  
targets, second row : gradient scaling, third row :  
adimensionalization.

dB and 10 dB, respectively. In particular, it is observed how the reduction of the control-space dimension reduces the fluctuations of the reconstructions near the boundary of the domain (especially for the reduced scattering coefficient maps). Results with a SNR equal to 20 dB are not presented because of high similarities with those obtained with a SNR of 30 dB.

The conclusion is that a mesh-based parameterization which consists in projecting onto a coarser mesh the variables involved in eq. (2.41), has the effect of regularizing implicitly the inverse problem by cutting off the high frequencies inherent to the gradient fluctuations due to noisy data. However, a too coarse mesh must not be chosen because such mesh does not allow a reconstruction that follows finely enough the fluctuations of the properties within the medium.



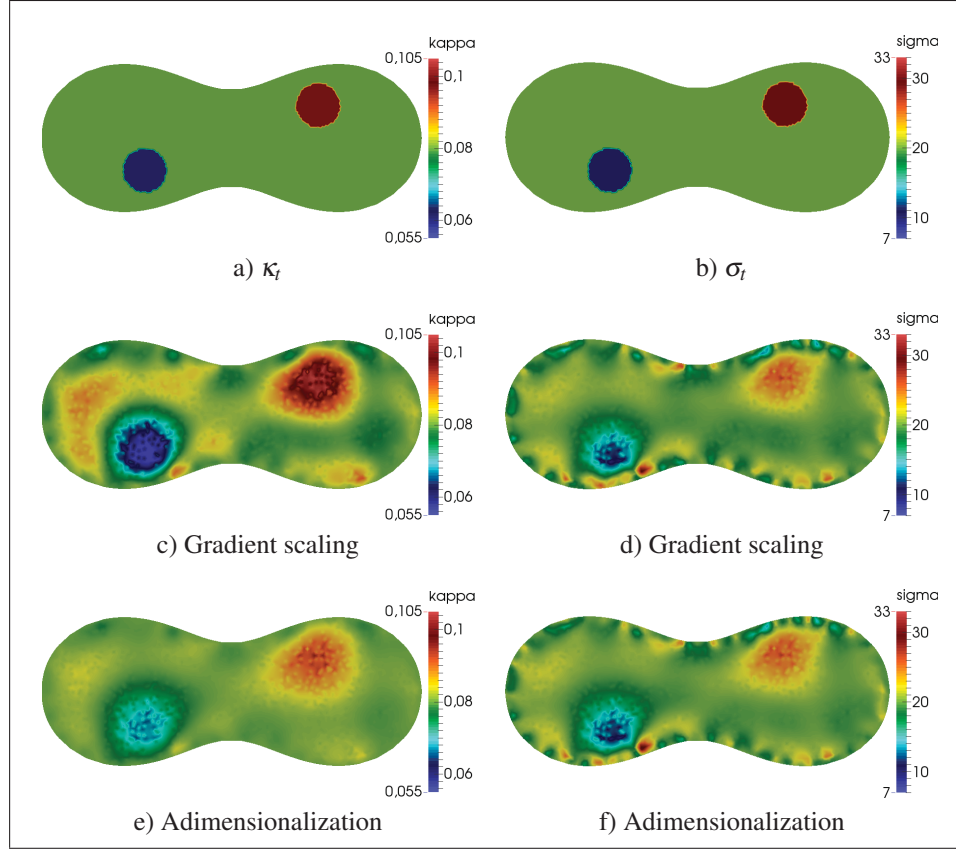


FIG. 2.6 BFGS algorithm ; mesh  $\mathcal{M}_1$  ; 10 dB noise.  
Reconstructions for both the gradient scaling approach and the  
adimensionalisation approach. Left :  $\kappa$ , right :  $\sigma$ . First row :  
targets, second row : gradient scaling, third row :  
adimensionalization.

### 2.5.6 BFGS algorithm : mixing mesh-based parameterization and Sobolev gradients method

The next step consists in using the previous result, i.e. use an appropriate parameterization of the control-space, and combine this first regularization tool with the Sobolev gradients method (Protas *et al.*, 2004). This regularization consists in choosing an appropriate inner product when extracting the cost function gradient, eq. (2.21). Most often, the  $L^2$ -inner product is that used to extract the gradient but, choosing another inner product such as the weighted Sobolev one can yield to much smoother and much closer-to-the-target reconstructions.

TAB. 2.3 BFGS algorithm. Obtained reconstruction errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for 4 meshes and 3 levels of noise.

Mesh \ Noise	10 dB	20 dB	30 dB
$\mathcal{M}_1$	$\begin{pmatrix} 4.22 & 8.95 \\ 0.73 & 0.69 \end{pmatrix}$	$\begin{pmatrix} 3.57 & 6.76 \\ 0.82 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 3.56 & 6.63 \\ 0.83 & 0.84 \end{pmatrix}$
$\mathcal{M}_2$	$\begin{pmatrix} 4.00 & 8.63 \\ 0.77 & 0.73 \end{pmatrix}$	$\begin{pmatrix} 3.46 & 6.28 \\ 0.84 & 0.86 \end{pmatrix}$	$\begin{pmatrix} 3.26 & 5.50 \\ 0.86 & 0.90 \end{pmatrix}$
$\mathcal{M}_3$	$\begin{pmatrix} 3.90 & 8.56 \\ 0.78 & 0.74 \end{pmatrix}$	$\begin{pmatrix} 3.32 & 5.75 \\ 0.85 & 0.88 \end{pmatrix}$	$\begin{pmatrix} 3.19 & 5.30 \\ 0.86 & 0.90 \end{pmatrix}$
$\mathcal{M}_4$	$\begin{pmatrix} 4.18 & 9.63 \\ 0.75 & 0.71 \end{pmatrix}$	$\begin{pmatrix} 3.45 & 6.00 \\ 0.84 & 0.87 \end{pmatrix}$	$\begin{pmatrix} 3.45 & 5.81 \\ 0.83 & 0.88 \end{pmatrix}$

Specifically, the inner product that is used here is :

$$(f, g)_{H^{1, \vartheta^S}(\Omega)} = (f, g)_{L^2(\Omega)} + \left( \vartheta^S \nabla f, \nabla g \right)_{L^2(\Omega)} \quad \forall f, g \in H^1(\Omega) \quad (2.47)$$

where  $\vartheta^S \in \mathbb{R}_+$  plays the role of regularization parameter and  $H^1(\Omega)$  is the Sobolev functional space whose definition can be easily deduced from the definition of  $\hat{H}^1(\Omega)$ , eq. (2.3), with  $f : \Omega \rightarrow \mathbb{R}$ .

This implies that :

$$\begin{aligned} \left( \nabla_{\mathcal{K}}^{H^{1, \vartheta^S}} j, \eta_{\mathcal{K}} \right)_{H^{1, \vartheta^S}(\Omega)} &= \left( \text{Re} \left( \kappa_{ap} \varphi \bar{\varphi}^* - n_{\Omega} D^2 \kappa_{ap} \nabla \varphi \cdot \nabla \bar{\varphi}^* \right), \eta_{\mathcal{K}} \right)_{\mathcal{Y}} \quad \forall \eta_{\mathcal{K}} \in H^1(\Omega) \\ \left( \nabla_{\zeta}^{H^{1, \vartheta^S}} j, \eta_{\zeta} \right)_{H^{1, \vartheta^S}(\Omega)} &= \left( \text{Re} \left( -n_{\Omega} D^2 \sigma_{ap} \nabla \varphi \cdot \nabla \bar{\varphi}^* \right), \eta_{\zeta} \right)_{\mathcal{Y}} \quad \forall \eta_{\zeta} \in H^1(\Omega) \end{aligned} \quad (2.48)$$

Thus, the use of such inner product when considering the cost function gradient extraction leads, after integration by parts, to :

$$\begin{aligned} \left( 1 - \vartheta^S \Delta \right) \nabla_{\mathcal{K}}^{H^{1, \vartheta^S}} j(\gamma) &= \sum_{l=1}^K \text{Re} \left( \kappa_{ap} \varphi_l \bar{\varphi}_l^* - n_{\Omega} D^2 \kappa_{ap} \nabla \varphi_l \cdot \nabla \bar{\varphi}_l^* \right) \quad \forall \mathbf{r} \in \Omega \\ \nabla \nabla_{\mathcal{K}}^{H^{1, \vartheta^S}} j(\gamma) \cdot \mathbf{n} &= 0 \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (2.49)$$

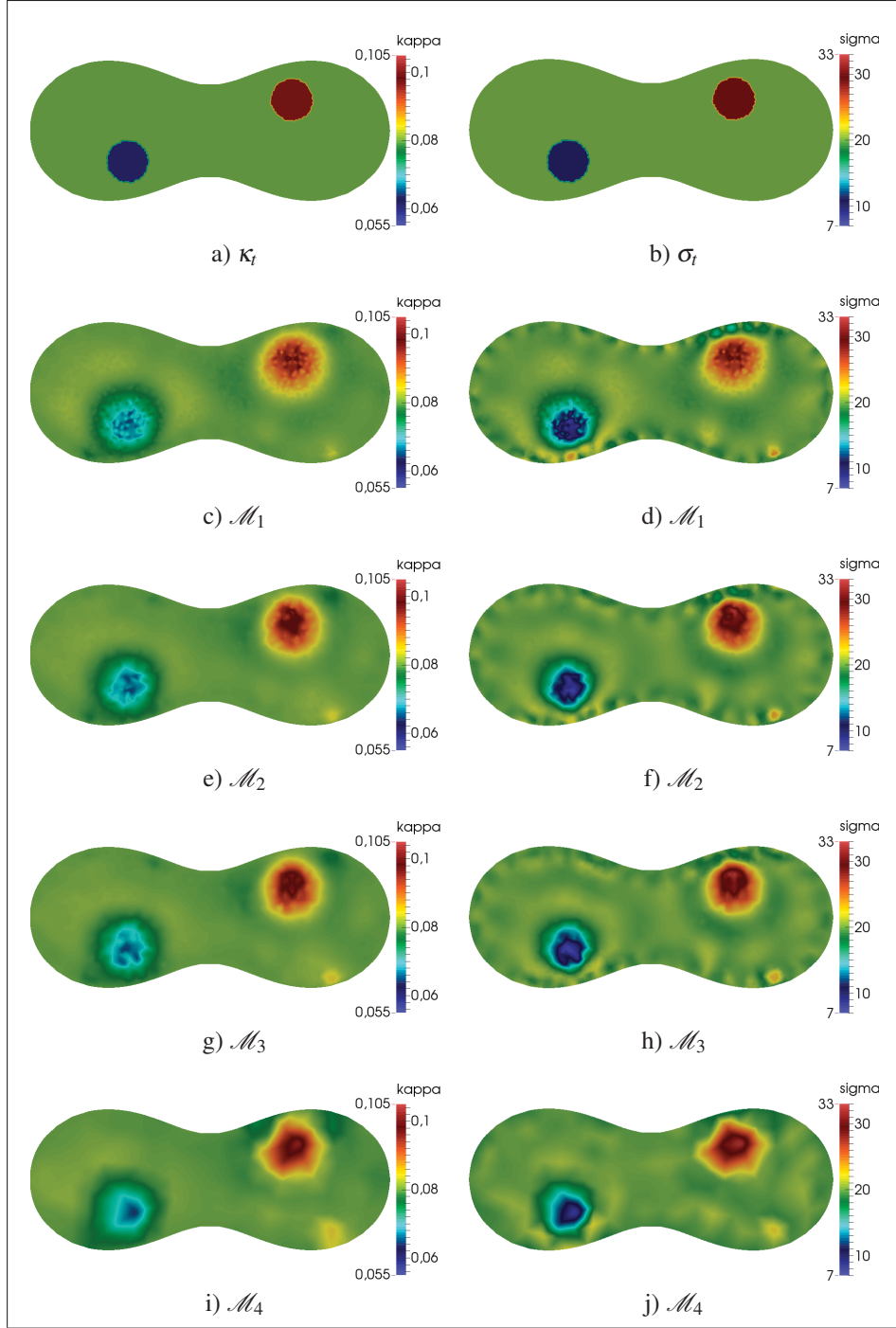


FIG. 2.7 BFGS algorithm ; 30 dB noise. Reconstructions for 4 control-space parameterizations. Left :  $\kappa$ , right :  $\sigma$ . First row : targets, next rows : reconstructions.

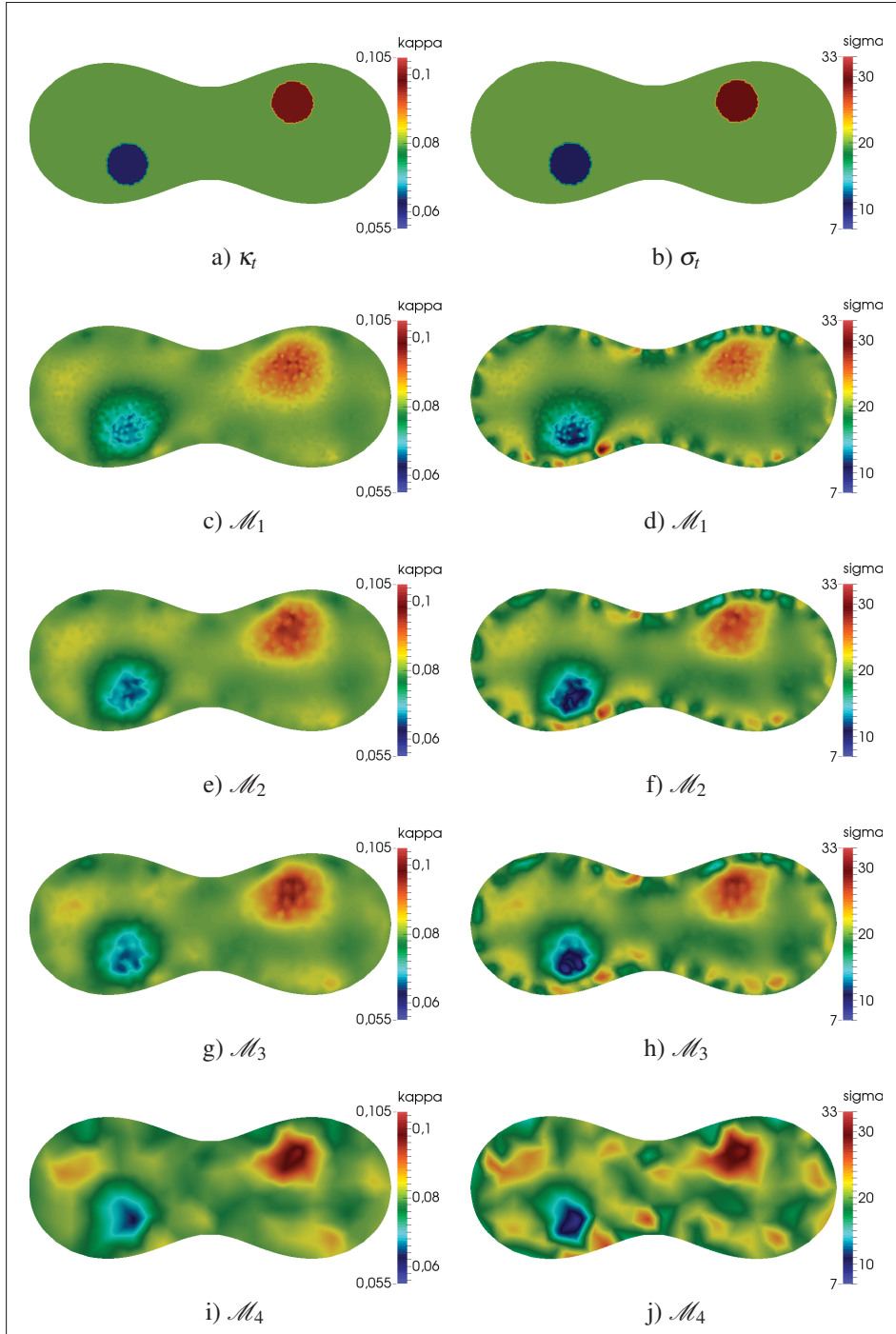


FIG. 2.8 BFGS algorithm ; 10 dB noise. Reconstructions for 4 control-space parameterizations. Left :  $\kappa$ , right :  $\sigma$ . First row : targets, next rows : reconstructions.

and

$$\begin{aligned} \left(1 - \vartheta^S \Delta\right) \nabla_{\zeta}^{H^1, \vartheta^S} j(\gamma) &= \sum_{l=1}^K \operatorname{Re} \left( -n_{\Omega} D^2 \sigma_{ap} \nabla \varphi_l \cdot \nabla \bar{\varphi}_l^* \right) \quad \forall \mathbf{r} \in \Omega \\ \nabla \nabla_{\zeta}^{H^1, \vartheta^S} j(\gamma) \cdot \mathbf{n} &= 0 \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (2.50)$$

Table 2.4 gives errors of some reconstructions with the use of such Sobolev gradients combined with appropriate parameterizations of the control-space for all the three SNR considered. Different values for the Sobolev regularization parameter were used, from  $\vartheta^S = 10^{-6}$  to  $\vartheta^S = 5 \times 10^{-2}$ . It is observed that errors obtained with  $\vartheta^S = 10^{-6}$  are equal or very close to the errors obtained when the ordinary  $L^2$ -inner product was used to extract the gradient. In this case, the value of the Sobolev parameter is too small and it is like no regularization is performed. Moreover, it is seen that the value  $\vartheta^S = 10^{-4}$  provides the best results for a SNR of 30 dB and the value  $\vartheta^S = 2 \times 10^{-3}$  provides the best results for a SNR of 20 dB. For these values of the Sobolev parameter, errors  $e_1$  ( $e_2$ ) are slightly smaller (larger) than those obtained when the  $L^2$ -inner product is used to extract the gradient. As expected, the optimal Sobolev parameter for a SNR equal to 30 dB is smaller than the one for a SNR equal to 20 dB, based on the simple fact that a low level of noise leads to a low need of regularization.

For a SNR equal to 10 dB, it is observed that no value of the Sobolev parameter leads to better results for all of the four errors  $e_1^K$ ,  $e_1^\sigma$ ,  $e_2^K$  and  $e_2^\sigma$ . More specifically, the best reconstruction for the absorption coefficient is obtained for  $\vartheta^S = 2 \times 10^{-3}$  while the best reconstruction for the reduced scattering coefficient is obtained for  $\vartheta^S = 5 \times 10^{-2}$ , which provides the worst result for the absorption coefficient.

Figures 2.9 and 2.10 show the final reconstructions for a SNR equal to 30 dB and 10 dB, respectively, and for all of the considered Sobolev parameters. In particular, it is observed that for  $\vartheta^S = 5 \times 10^{-2}$  (Figs. 2.9-i, 2.9-j, 2.10-i, 2.10-j), the reconstructions present very few fluctuations compared to those obtained with smaller values of the Sobolev parameter. Actually, this illustrates how the Sobolev gradients act as a smoother so that the high-frequency fluctuations present in the adjoint variable are removed or at least de-emphasized through the cost function computation.

To go one step further and attempt to improve the quality of the reconstructions in presence of a high level of noise such as with a SNR equal to 10 dB, the next subsection investigates the choice of a space-dependent Sobolev parameter.

TAB. 2.4 BFGS algorithm ; mesh  $\mathcal{M}_3$ . Obtained reconstruction errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for 4 values for the Sobolev regularization parameter. 3 levels of noise are considered.

$\vartheta^S \setminus \text{Noise}$	10 dB	20 dB	30 dB
$1 \times 10^{-6}$	$\begin{pmatrix} 3.90 & 8.56 \\ 0.78 & 0.74 \end{pmatrix}$	$\begin{pmatrix} 3.32 & 5.75 \\ 0.85 & 0.88 \end{pmatrix}$	$\begin{pmatrix} 3.21 & 5.32 \\ 0.86 & 0.90 \end{pmatrix}$
$1 \times 10^{-4}$	$\begin{pmatrix} 3.90 & 8.61 \\ 0.78 & 0.74 \end{pmatrix}$	$\begin{pmatrix} 3.33 & 5.80 \\ 0.85 & 0.88 \end{pmatrix}$	$\begin{pmatrix} 3.19 & 5.28 \\ 0.86 & 0.90 \end{pmatrix}$
$2 \times 10^{-3}$	$\begin{pmatrix} 3.88 & 8.99 \\ 0.78 & 0.73 \end{pmatrix}$	$\begin{pmatrix} 3.31 & 5.67 \\ 0.85 & 0.89 \end{pmatrix}$	$\begin{pmatrix} 3.22 & 5.23 \\ 0.86 & 0.91 \end{pmatrix}$
$5 \times 10^{-2}$	$\begin{pmatrix} 4.05 & 7.78 \\ 0.77 & 0.78 \end{pmatrix}$	$\begin{pmatrix} 3.65 & 6.57 \\ 0.81 & 0.85 \end{pmatrix}$	$\begin{pmatrix} 3.64 & 6.53 \\ 0.81 & 0.85 \end{pmatrix}$

### 2.5.7 BFGS algorithm : space-dependent Sobolev gradients method

The key idea behind the development of a space-dependent Sobolev gradients method is quite simple. It has been observed that final reconstructions present undesirable fluctuations coming from the noisy data which propagate to the adjoint variable and then to the gradient. The Sobolev gradients method can lead to better results by smoothing the high-frequencies within the gradient. But, these fluctuations are mainly present near the edge of the computational domain while the Sobolev regularization is uniform within the medium. Hence, a method which strongly regularizes the gradient near the medium boundary but very weakly in the bulk of the medium seems to be an effective strategy to improve the quality of the reconstructions.

To do so, the real positive constant  $\vartheta^S$  involved in the inner product eq. (2.47) is replaced by the following real positive function :

$$\vartheta_v^S(\mathbf{r}) = \exp(-vr^2(\mathbf{r})) \quad (2.51)$$

where  $r(\mathbf{r})$  is a space-dependent function which gives the shortest distance from  $\mathbf{r}$  to the boundary  $\partial\Omega$  and  $\nu \in \mathbb{R}_+$  now plays the role of the regularization parameter. As an example, Fig. 2.11 depicts the function  $\vartheta_{25}^S(\mathbf{r})$ .

The partial differential equation allowing the computation of the cost function gradients when using such an inner product is given by :

$$\begin{aligned} \nabla_{\mathcal{Z}}^{H^1, \vartheta_{\nu}^S} j(\gamma) - \nabla \cdot \left( \vartheta_{\nu}^S \nabla \nabla_{\mathcal{Z}}^{H^1, \vartheta_{\nu}^S} j(\gamma) \right) &= \sum_{l=1}^K \operatorname{Re} \left( \kappa_{ap} \varphi_l \bar{\varphi}_l^* - n_{\Omega} D^2 \kappa_{ap} \nabla \varphi_l \cdot \nabla \bar{\varphi}_l^* \right) \quad \forall \mathbf{r} \in \Omega \\ \nabla \nabla_{\mathcal{Z}}^{H^1, \vartheta_{\nu}^S} j(\gamma) \cdot \mathbf{n} &= 0 \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (2.52)$$

and

$$\begin{aligned} \nabla_{\zeta}^{H^1, \vartheta_{\nu}^S} j(\gamma) - \nabla \cdot \left( \vartheta_{\nu}^S \nabla \nabla_{\zeta}^{H^1, \vartheta_{\nu}^S} j(\gamma) \right) &= \sum_{l=1}^K \operatorname{Re} \left( -n_{\Omega} D^2 \sigma_{ap} \nabla \varphi_l \cdot \nabla \bar{\varphi}_l^* \right) \quad \forall \mathbf{r} \in \Omega \\ \nabla \nabla_{\zeta}^{H^1, \vartheta_{\nu}^S} j(\gamma) \cdot \mathbf{n} &= 0 \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (2.53)$$

**Remark 10.** The limit  $\nu \rightarrow 0$  corresponds to the use of the space-independent Sobolev gradients method with  $\vartheta^S = 1$ , and the limit  $\nu \rightarrow +\infty$  corresponds to the use of the ordinary  $L^2$ -inner product when extracting the gradients.

The BFGS algorithm has been run with the use of the space-dependent Sobolev gradients method for values of the regularization parameter  $\nu$  ranging from 1 to 100. The errors are reported in Fig 2.12. The thick black line represents the errors obtained when the ordinary  $L^2$ -inner product was used to extract the gradient, and the points represent the errors for the proposed space-dependent Sobolev gradients. The points are set red for worse errors (when  $e_1|_{H^1, \vartheta_{\nu}^S} > e_1|_{L^2}$  or  $e_2|_{H^1, \vartheta_{\nu}^S} < e_2|_{L^2}$ ), and points are set green otherwise, for better errors (when  $e_1|_{H^1, \vartheta_{\nu}^S} < e_1|_{L^2}$  or  $e_2|_{H^1, \vartheta_{\nu}^S} > e_2|_{L^2}$ ).

In most cases, reconstructions are improved (the very few cases with very small  $\nu$  correspond to space-independent Sobolev gradients with  $\vartheta^S \rightarrow 1$  which is too high as shown in previous section). Indeed, as soon as  $\nu \gtrsim 10$ , both the deviation factor and the correlation coefficient are better for both the absorption function  $\kappa$  and the reduced scattering function  $\sigma$ . For  $15 \lesssim \nu \lesssim 30$ , reconstructions are indeed really improved, especially for the reduced scattering function  $\sigma$

whose deviation factor has been reduced about 16 % and correlation coefficient improved about 11 % when compared to the best result obtained when the ordinary  $L^2$ -inner product was used to extract the gradient. Figure 2.13 illustrates this result by comparing the final reconstructions obtained with the  $L^2$ -inner product and with the space-dependent Sobolev gradients method for  $\nu = 25$ . As expected, the latter reconstructions contain much less fluctuations, especially close to the boundary of the domain, and reconstructions are closer to the targets. Obtained errors for  $\nu = 25$  are the following : 
$$\begin{pmatrix} e_1^{\kappa} & e_1^{\sigma} \\ e_2^{\kappa} & e_2^{\sigma} \end{pmatrix} = \begin{pmatrix} 3.78 & 7.17 \\ 0.80 & 0.82 \end{pmatrix}.$$

## 2.6 Conclusion

Several numerical algorithms have been implemented for the solution of a two-dimensional DOT problem. On one side two distinct optimization algorithms have been considered : i) the damped GN algorithm ; and ii) the BFGS algorithm. On the other side, several regularization tools have been developed : a) mesh-based parameterization of the control variables ; b) Tikhonov regularization ; and c) Sobolev gradients method. The Sobolev gradients comes from the use of the weighted Sobolev inner product when extracting the cost function gradient. Although this is a new regularization in the field of the DOT, it has already been extended to a space-dependent version. The aim of this paper was also to compare the methods when combined together. More specifically, the mesh-based parameterization has been coupled with both the damped GN algorithm and the BFGS algorithm. The Tikhonov regularization has been coupled with the GN algorithm (it is almost ineffective coupled with the BFGS algorithm). Finally, both space-independent and space-dependent Sobolev gradients methods have been successfully coupled with the mesh-based parameterization of the control space and the BFGS optimization algorithm to enhance the reconstructions. Three levels of noise were considered and deviation factors as well as correlation coefficients were used to perform quantitative comparisons.

For the damped GN algorithm, the study coupling the mesh-based parameterization regularization with the ordinary Tikhonov regularization showed that the quasi-optimal Tikhonov



parameter is quasi-independent of the control-space parameterization in the under-determined case. Moreover, it has been observed that the damped Gauss-Newton algorithm is very few sensitive to the level of noise in the under-determined case whereas it is a little bit sensitive to the level of noise in the over-determined case. Also, it has been shown that, for a given level of noise, a reduction of the control-space dimension leads to improved results. Hence, the best reconstructions have been obtained in the over-determined case for all of the levels of noise considered.

For the BFGS algorithm, it has been first shown that the adimensionalization strategy provides better reconstructions than the scaling of the cost function gradients. Then, it has been observed that, unlike the damped Gauss-Newton algorithm, the BFGS algorithm is sensitive to the level of noise whatever the parameterization of the control variables : the lesser the level of noise, the better the reconstructions. Also, it has been shown that a reduction of the control-space dimension, until a certain limit, leads to more accurate final reconstructions. For all of the levels of noise considered, the same intermediate mesh has provided the best results. Finally, the use of the Sobolev gradients method has been investigated. It was found that, excepted for the highly noisy case, some values of the Sobolev parameter lead to slightly more accurate reconstructions for both the absorption and reduced scattering coefficients. Moreover, the larger the level of noise, the larger the optimal Sobolev parameter. Finally, for the first time to the best knowledge of the authors, a space-dependent Sobolev gradients method was developed. This method was successfully applied to the highly noisy case (10 dB) : a significant improvement of the reconstruction accuracies was observed for a large range of the space-dependent Sobolev regularization parameter. Although results of this latter method in the case of small levels of noise were not given in order not to overburden this paper, these improvements compared to the space-independent Sobolev gradients method also hold in these cases.

Based on the numerical results obtained for the proposed application, one can conclude that the BFGS algorithm outperforms the damped Gauss-Newton in several aspects. First, it was found that the computational cost per iteration is lesser for the BFGS algorithm than for the damped Gauss-Newton algorithm. Second, one can remark that, in contrast to the BFGS algorithm, the

damped Gauss-Newton algorithm requires some penalization of the cost function associated to the difficult and/or CPU-time consuming determination of an appropriate penalization parameter in order to be stable. Third, it has been observed that the BFGS algorithm associated to an appropriate parameterization of the control variables provides more accurate reconstructions than the damped Gauss-Newton algorithm associated to an appropriate parameterization of the control variables combined with the Tikhonov regularization for all the levels of noise considered and for both the radiative properties, excepted for the reduced scattering coefficient for the 10 dB case. But, adding the space-dependent Sobolev gradients regularization to the BFGS algorithm, results become far better than those obtained with the damped Gauss-Newton algorithm for the 10 dB case. Moreover, it should be noted the ease of finding an appropriate value for the parameter involved in the space-dependent Sobolev gradients regularization method.

Upcoming work include theoretical developments about the quasi-independence of the quasi-optimal Tikhonov parameter with the parameterization of the control variables, the extension of the proposed regularized BFGS algorithm to the three-dimensional case and the study on the choice of the space-dependent positive function employed in the space-dependent Sobolev gradients method. The application of the proposed method to the inversion of the full radiative transfer equation is also considered.

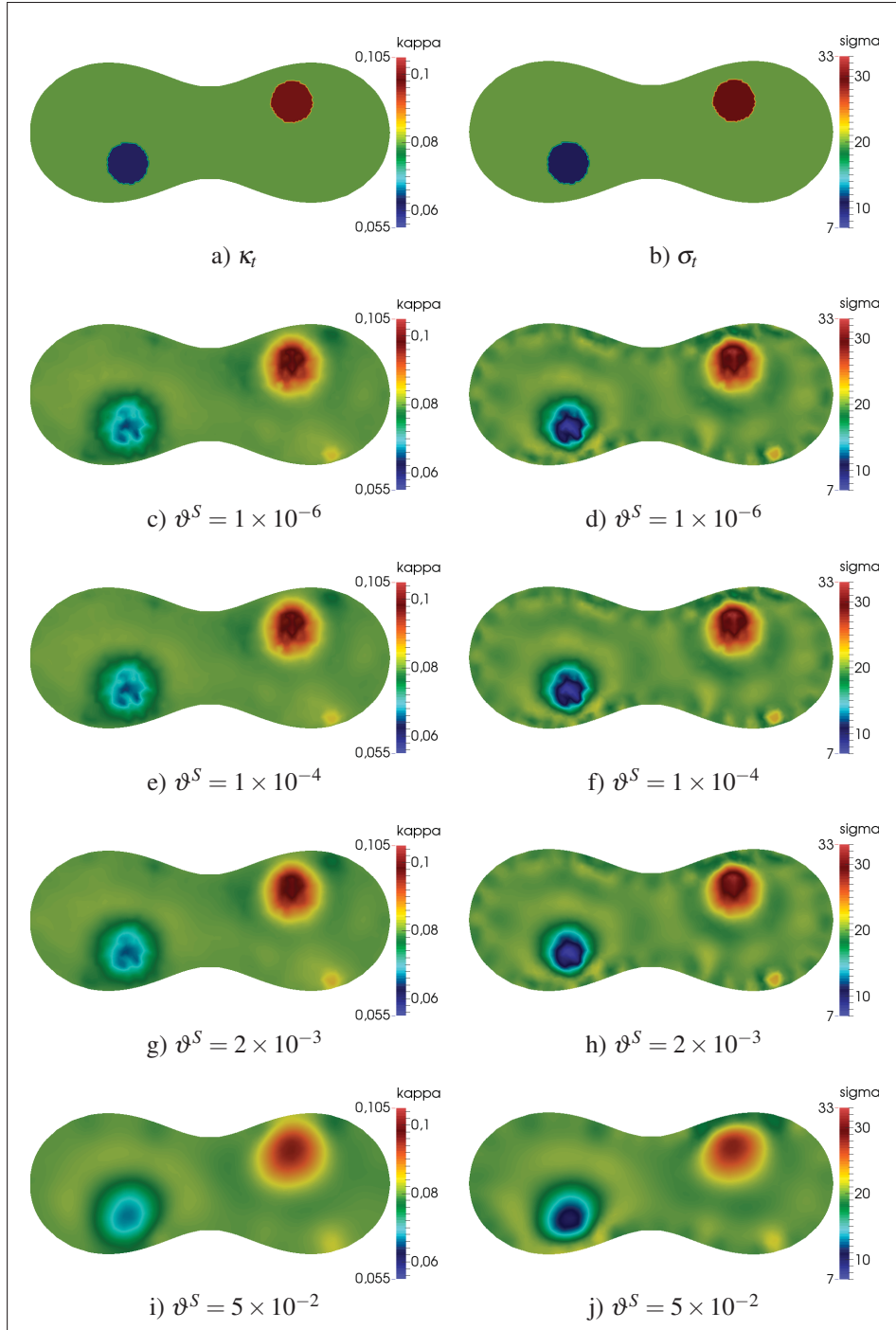


FIG. 2.9 BFGS algorithm ; mesh  $\mathcal{M}_3$ , 30 dB noise. Obtained reconstruction errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for 4 values for the Sobolev regularization parameter. Left :  $\kappa$ , right :  $\sigma$ . First row : targets, next rows : reconstructions.

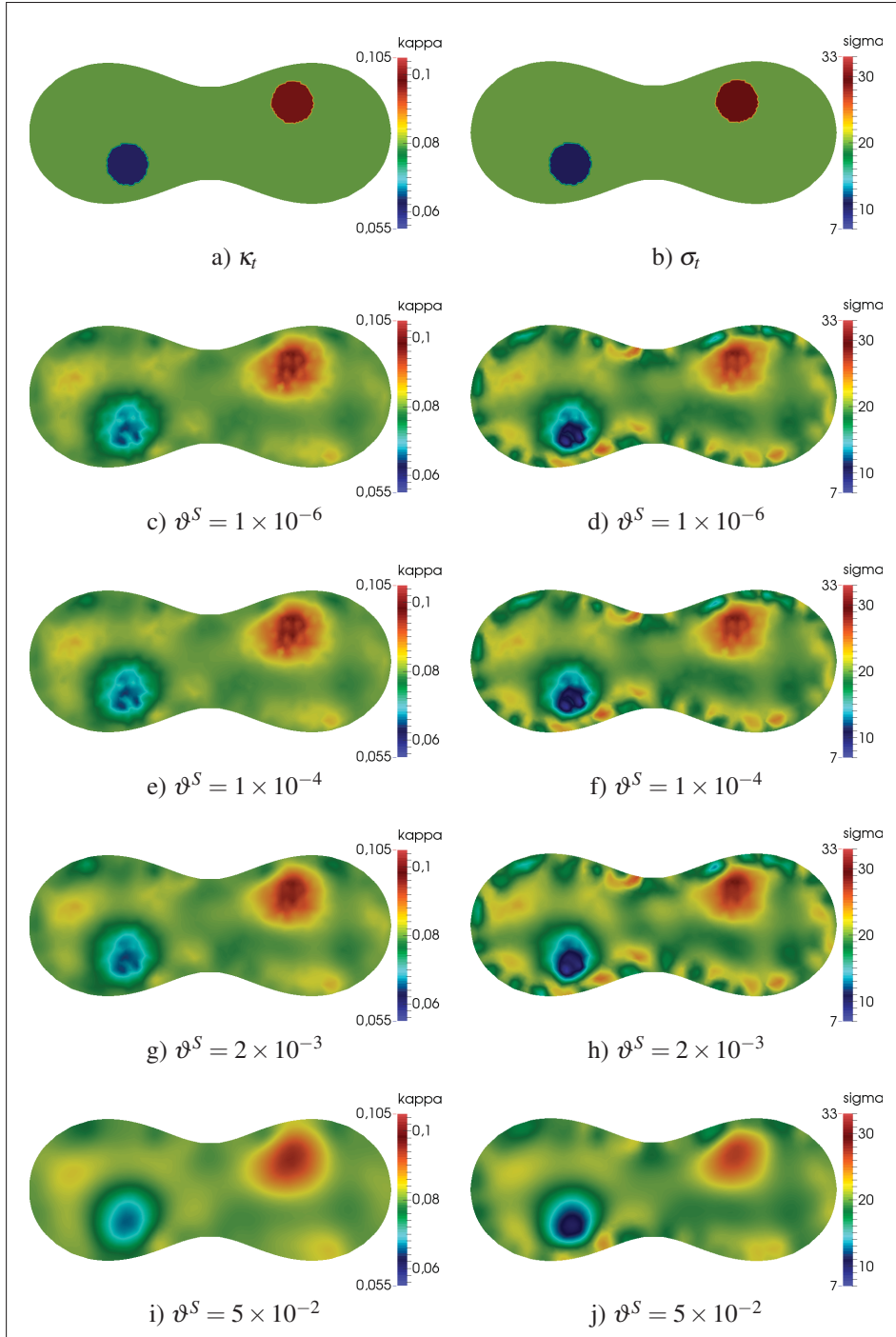


FIG. 2.10 BFGS algorithm ; mesh  $\mathcal{M}_3$ , 10 dB noise. Obtained reconstruction errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for 4 values for the Sobolev regularization parameter. Left :  $\kappa$ , right :  $\sigma$ . First row : targets, next rows : reconstructions.

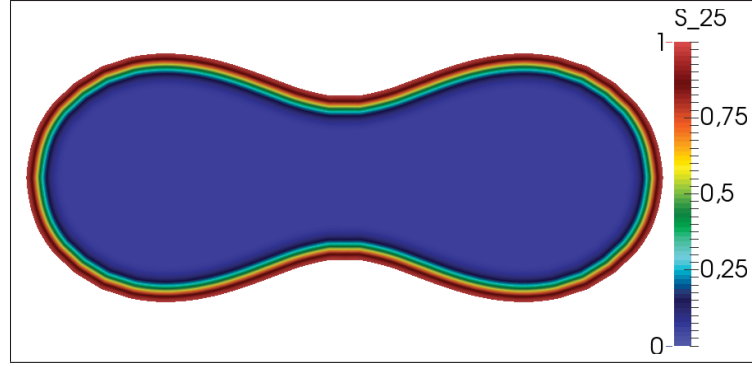


FIG. 2.11 The space-dependent Sobolev filter  $\vartheta_{25}^S(\mathbf{r})$ .

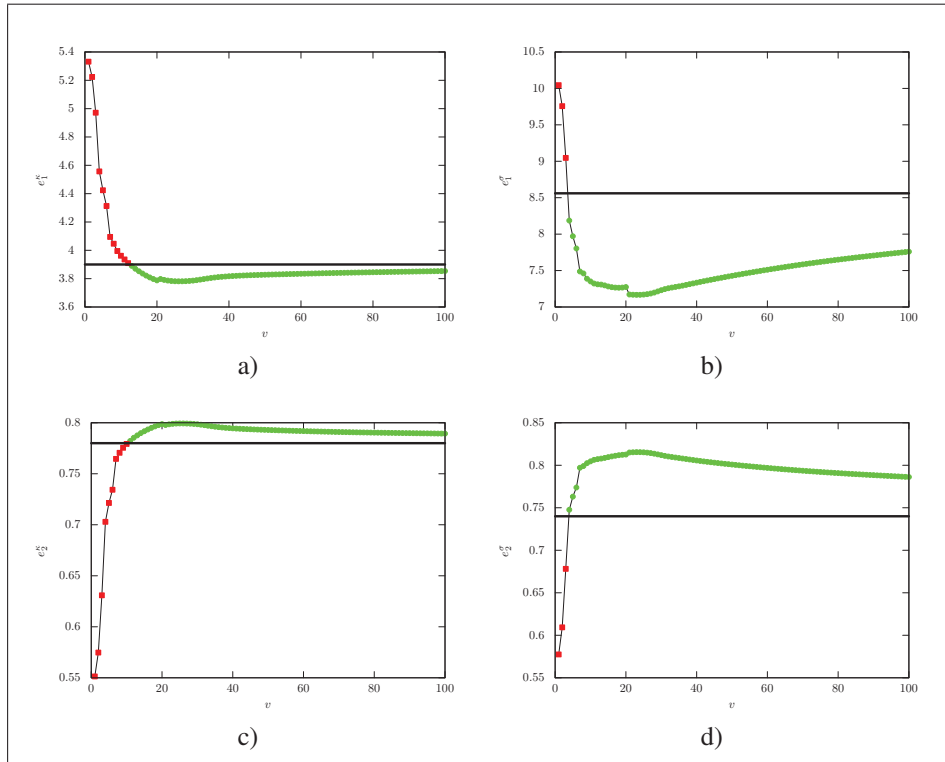


FIG. 2.12 Red and green points : evolution of errors  $e_1^K$  (a),  $e_1^\sigma$  (b),  $e_2^K$  (c) and  $e_2^\sigma$  (d) with respect to the regularization parameter  $\nu$ . The parameter  $\nu$  ranges from 1 to 100. Black line : errors  $e_1^K$ ,  $e_1^\sigma$ ,  $e_2^K$  and  $e_2^\sigma$  found in Table 2.3 (i.e. obtained with the BFGS method), and for 10 dB and the mesh  $\mathcal{M}_3$  (i.e. corresponding to the best control-space parameterization).

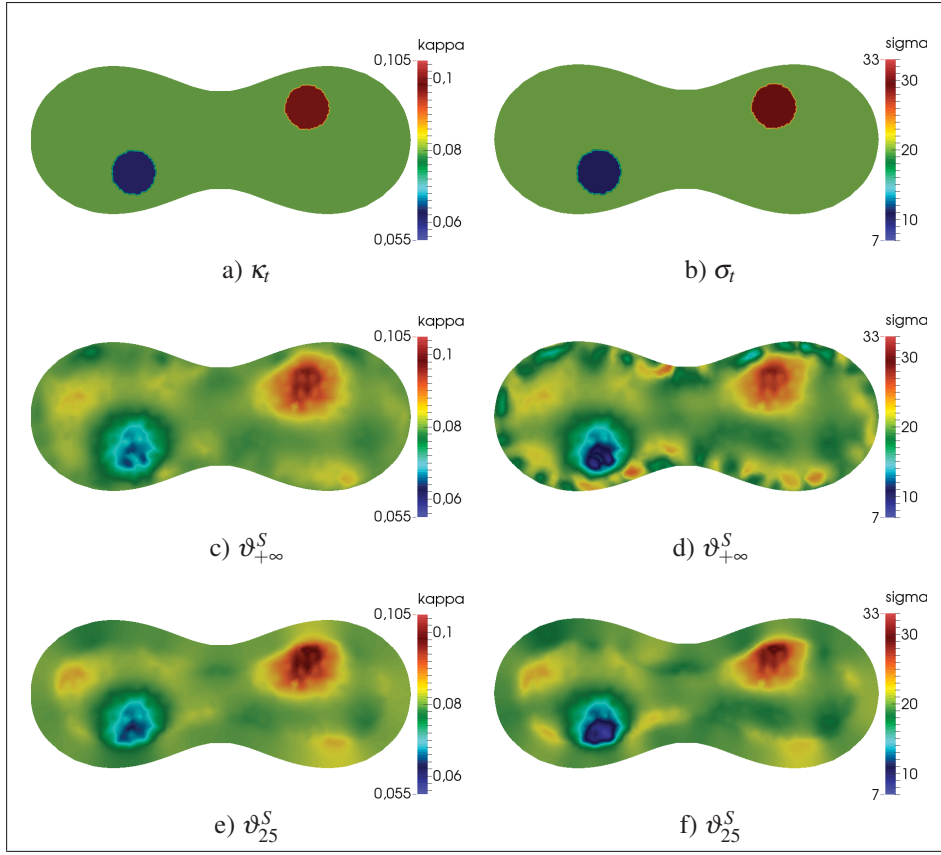


FIG. 2.13 BFGS algorithm ; mesh  $\mathcal{M}_3$ , 10 dB noise.  
 Reconstructions with the space-dependent Sobolev regularization  
 with  $\nu = 25$ . Left :  $\kappa$ , right :  $\sigma$ . First raw : targets, second raw :  
 reconstructions with the ordinary  $L^2$ -inner product, third raw :  
 reconstructions with the  $\vartheta_{25}^S(\mathbf{r})$  Sobolev filter.

## CHAPITRE 3

### ARTICLE 2 : QUASI-OPTIMAL TIKHONOV PENALIZATION AND PARAMETERIZATION COARSENESS IN SPACE-DEPENDENT FUNCTION ESTIMATION

Fabien Dubot<sup>1,2</sup>, Yann Favennec<sup>2</sup>, Benoit Rousseau<sup>2</sup>, Yvon Jarny<sup>2</sup>, Daniel Rousse<sup>1</sup>

<sup>1</sup> Chaire de recherche industrielle t3e, École de technologie supérieure,  
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

<sup>2</sup> Laboratoire de thermocinétique de Nantes, UMR CNRS 6607,  
Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France

Article accepté dans la revue « Inverse Problems in Science and Engineering » en avril 2015.

#### Abstract

The determination of space-dependent functions from boundary measurements or inner pointwise measurements are ill-posed inverse problems that require regularization tools to be stabilized. Among numerous regularization strategies, the Tikhonov penalization is one of the most used in the field of space-dependent function estimation. Its efficient use relies on the Tikhonov parameter value for which search is time consuming although necessary, specially in the field of nonlinear inversion. Other strategies, such as appropriate parameterization, have recently proven to be very efficient to cope with the ill-posedness of such problems. This paper shows that the optimal Tikhonov parameter is almost independent of the mesh used to project the functions to be retrieved. As a consequence, this value should be sought using a coarse mesh even though reconstructions could further be done on finer meshes. This conclusion is validated by numerical means.

#### 3.1 Introduction

The reconstruction of space-dependent functions from pointwise measurements belongs to inverse problems known to be difficult to be solved due to their ill-posed character. More spe-

cifically, the reconstruction of maps of radiative properties by illuminating a semi-transparent medium with near infrared radiation and measuring emerging radiation is by definition the classical diffuse optical tomography (DOT) inverse problem.

From a mathematical point of view, within the diffuse approximation (DA) framework in the frequency domain (Arridge, 1999), the geometry ( $\Omega$ ) being fixed, the total inward flux ( $f_\omega$ ) of the amplitude modulated diffuse source at the frequency  $\omega$  being known as well as the speed of light ( $c$ ) in the medium, the knowledge of the space-dependent absorption ( $\kappa$ ) and reduced scattering ( $\sigma$ ) coefficients is sufficient to simulate the photon density distribution ( $\varphi$ ) in each location of the computational domain. Such modeling is the forward problem which is known to be well-posed mathematically.

Contrarily, the reconstruction of physical properties from the knowledge of photon density on sensors constitutes an inverse problem. Unfortunately, such a problem is ill-posed in the sense of Hadamard (Hadamard, 1902). Hence, regularization must be used to stabilize the solution. Several regularization tools have been introduced in the field of DOT these last decades, among which, the use of appropriate control space parameterization (Gu *et al.*, 2003; Chavent, 2010; Favennec *et al.*, 2013; Balima *et al.*, 2013). Dealing with matrix-based inversion, the Tikhonov regularization method, which is a popular method, consists in adding a penalization term to the cost function to be minimized ; this method actually relies on a weight parameter that is to be determined very carefully since : (i) under-regularization leads to small cost functions at the end but at the price of highly fluctuating property maps ; and (ii) over-regularization leads to stable property maps around priors, leading to biased solutions (Morozov *et al.*, 1984).

In the following, the optimal Tikhonov parameter,  $\vartheta^*$ , is defined as the one that minimizes the distance between the actual solution and the noise-free solution. This distance can be computed with  $E(\vartheta) = \frac{\|\gamma_\vartheta - \bar{\gamma}\|}{\|\bar{\gamma}\|}$ , where  $\vartheta$  is the Tikhonov parameter,  $\gamma_\vartheta$  is the solution given for the Tikhonov parameter  $\vartheta$ ,  $\bar{\gamma}$  is the exact solution, and  $\|\cdot\|$  is a norm. The point is that one cannot, in practice, compute the optimal Tikhonov parameter,  $\bar{\gamma}$  being unknown. However, numerous heuristic methods have been introduced to compute quasi-optimal Tikhonov parameters.



Some methods use the assumed to be known standard deviation such as in the discrepancy principle (Morozov *et al.*, 1984; Goncharskii *et al.*, 1973). Others (Golub *et al.*, 1979; Hansen et O’Leary, 1993; O’Leary, 2001) use less information about the noise properties, such as for instance the Generalized Cross Validation (GCV), or even no information about the noise level present in the system, such as the well known and controversial L-curve. The methods used to find the optimal Tikhonov parameter or at least a quasi-optimal parameter is beyond of the scope of this paper. Moreover, even though methods that do not use error informations are widely used for solving practical problems in various engineering areas, let us recall that in general the L-curve method for instance, because it does not rely on any noise information, is, as shown by (Yagola *et al.*, 2002; Vogel, 1996), not convergent and introduces a nonremovable bias.

In this paper, the diffuse optical tomography (DOT) problem is formulated as a nonlinear least squares problem and solved by the Gauss-Newton (GN) method. Though this method has been widely used these last years to solve this inverse problem (Paulsen et Jiang, 1995; Schweiger *et al.*, 2005; Niu *et al.*, 2008; Tarvainen *et al.*, 2008a; Dehghani *et al.*, 2009a,b), it is worth mentioning that the GN method is only locally convergent, and this is so, only when the initial guess is not too far from the minimum and the inverse problem is mildly nonlinear (Björck, 1996). This means that some choices of the initial guess far from the solution lead to inaccurate results whatever the regularization method used. To illustrate this property, the dependence of the accuracy of the numerical solution with respect to the initial guess is discussed in the numerical results section. Although the study of the initial guess of the solution on the convergence of the optimizer is not the main issue of this paper, it should be noted that globally convergent methods, which overcome the major drawback of the GN optimizer, have been developed and applied to the DOT problem (Su *et al.*, 2006; Shan *et al.*, 2008a,b; Beilina et Klibanov, 2008, 2012; Su *et al.*, 2014).

The remainder of the paper will show that the optimal Tikhonov parameter is almost independent of the mesh coarsening used for the parameterization. This assertion is demonstrated theoretically and verified numerically based on two different under-determined inverse pro-

blems of space-dependent function estimation : a linear inverse heat conduction problem of flux estimation and a nonlinear DOT problem. The studies are performed on synthetic data and therefore the errors on reconstructions could be calculated. For this reason one can somehow consider that the proposed approach is heuristic.

Before dealing with the nonlinear inverse problem of DOT (which is the main objective of this paper), a steady-state two-dimensional inverse heat conduction problem of space-dependent heat flux estimation is dealt with in section 3.2. Such inverse problem has recently been studied (Hensel et Hills, 1989; Taler, 1992; Dulikravich et Martin, 1996; Jarny, 2011). This space-dependent heat flux estimation problem is highly different from the DOT problem because : (i) it deals with a different physics : heat conduction ; (ii) measurements are performed within the domain rather than on a boundary ; (iii) the heat flux on a boundary is the function to be retrieved rather than physical properties within the medium ; and (iv) the forward problem is linear with respect to the coefficients involved in the space-dependent finite element heat flux parameterization. The interest of this first study is above all to illustrate that the quasi-independence of the optimal Tikhonov parameter with the control space parameterization is found in a less heuristic way since only measurement errors are used in the whole process. Indeed, the determination of the Tikhonov parameter can be performed straightforwardly through the knowledge of measurement errors only. In this case, the discrepancy principle is the major recipe to compute the parameters involved in the space-dependent heat flux parameterization. It is known that this method gives satisfactory results and other methods found in literature may give a Tikhonov parameter closer to the optimal one (Golub *et al.*, 1979; Hansen et O’Leary, 1993; O’Leary, 2001) but such consideration is out of the scope of this paper. Moreover, the statement of this article is also validated on this linear inverse problem in the sense of the optimal Tikhonov parameter  $\vartheta^*$  previously defined as  $\arg \min E(\vartheta)$ .

Section 3.3 deals with the nonlinear DOT problem. The forward model of light propagation in a highly diffuse medium is first presented along with the inverse problem settings. First and second order cost function derivatives used afterwards in optimization are then detailed. The finite element parameterization of, on one side, the state and its derivative and, on the other

side, the properties to be retrieved, are presented. The optimization is then written down based on the proposed finite dimensional control space. This section then presents the most usual Tikhonov regularization and proves that, under some weak hypothesis, the optimal Tikhonov parameter is quasi-independent of the dimension of the control space. A numerical verification is performed on the nonlinear DOT problem based on the theoretical demonstration.

Overall, the conclusion is that, when dealing with nonlinear inverse problems demanding heavy computational time, the determination of the Tikhonov parameter, whatever the method chosen among those of (Tikhonov *et al.*, 1998) for instance, should be preferably performed on a highly coarse mesh in order to lower computational time.

## 3.2 Space-dependent heat flux estimation

### 3.2.1 Problem statement

In this first study, a linear steady-state two-dimensional inverse heat conduction problem is considered to illustrate the quasi-independence of the Tikhonov parameter with mesh coarsening used for the finite element projection of the quantity of interest. In the present case already studied in different contexts (Hensel et Hills, 1989; Taler, 1992; Dulikravich et Martin, 1996; Jarny, 2011), a space-dependent heat flux on a part of the boundary is to be retrieved from pointwise measurements within a bounded domain  $\Omega = [0, 1]^2$ . This application is treated first because of its simplicity : as a matter of fact, the response being linear with respect to the input flux (Jarny, 2011), the cost function to be minimized is purely quadratic, and specific tools such as the singular value decomposition coupled with the maximum discrepancy principle (Morozov *et al.*, 1984; Goncharskii *et al.*, 1973) can be used to compute straightforwardly a quasi-optimal Tikhonov parameter. In this case, one can speak of “Tikhonov parameter found according to the discrepancy principle”. It is shown elsewhere (O’Leary, 2001) that this quasi-optimal Tikhonov parameter is likely to be close to the optimal one.

Homogeneous steady-state heat conduction without source term but with mixed boundary conditions is described such that :

$$\begin{cases} -\Delta T = 0 & (r_1, r_2) \in \Omega = ]0, 1[ \times ]0, 1[ \\ T = 0 & r_1 = 0 \\ \nabla T \cdot \mathbf{n} = 0 & r_2 = 0 \text{ and } r_2 = 1 \\ -\lambda_c \nabla T \cdot \mathbf{n} = \varphi & r_1 = 1 \end{cases} \quad (3.1)$$

where  $T$  is the temperature,  $\lambda_c$  is the thermal conductivity and  $\varphi$  is a space-dependent heat flux. The forward problem consists in solving eq. (3.1) for  $T$  assuming  $\lambda_c$  and  $\varphi(r_2)$  are known. In contrast, the inverse problem consists in estimating the space-dependent heat flux  $\varphi(r_2)$  on the basis of suitable measurement data  $\check{T}^l$  in  $\Omega$ , under the assumption that the value of  $\lambda_c$  is known.

Let us consider  $k$  inner pointwise measurements  $(r_1^l, r_2^l)$ ,  $l = 1, \dots, k$ . Discrepancies between predictions  $T^l = T(r_1^l, r_2^l)$  and associated measurements  $\check{T}^l$ ,  $l = 1, \dots, k$ , are integrated to the cost function :

$$j_{\vartheta}(\varphi) = \mathcal{J}(T) + \mathcal{J}^+(\varphi) = \|T - \check{T}\|_{\mathbb{R}^k}^2 + \vartheta \|\phi\|_{\mathbb{R}^{\Xi}}^2 \quad (3.2)$$

with  $\|a\|_{\mathbb{R}^k}^2 = \sum_{l=1}^k a_l^2$  and  $\|a\|_{\mathbb{R}^{\Xi}}^2 = \sum_{l=1}^{\Xi} a_l^2$ .  $\vartheta$  is the Tikhonov parameter to be searched according to the noise level  $\varepsilon$ .  $\phi$  is the vector obtained after finite element parameterization of the heat flux  $\varphi$ , that is  $\varphi(r_2) = \sum_{l=1}^{\Xi} \Theta_l(r_2) \varphi(r_{2l}) = \sum_{l=1}^{\Xi} \Theta_l(r_2) \phi_l$  where  $(\Theta_l)_{l=1}^{\Xi}$  is a finite element basis of  $[0, 1]$ . The singular value decomposition of the matrix  $A : \phi \in \mathbb{R}^{\Xi} \mapsto T \in \mathbb{R}^k$  such that  $A = W \Lambda V^t$ , enables us to rewrite the cost function such that :

$$j_{\vartheta}(\xi) = \left\| \Lambda \xi - \check{\xi} \right\|_{\mathbb{R}^k}^2 + \vartheta \|\xi\|_{\mathbb{R}^{\Xi}}^2 \quad (3.3)$$

where  $\xi = V^t \phi$  and  $\check{\xi} = W^t \check{T}$ . With such a decomposition, the solution of  $\xi_{\vartheta}^* = \arg \min j_{\vartheta}(\xi)$  is given in terms of singular values of  $A$ ,  $(\eta_i)_{i=1}^{\Xi}$  :

$$\xi_{\vartheta i}^* = \frac{\eta_i \check{\xi}_i}{\eta_i^2 + \vartheta} \quad \forall i = 1, \dots, \Xi \quad (3.4)$$

On the other hand, the discrepancy principle leads to determine  $\xi_{\vartheta}^*$  such that :

$$\left\| \Lambda \xi_{\vartheta}^* - \check{\xi} \right\|_{\mathbb{R}^k}^2 = \varepsilon^2 \quad (3.5)$$

where  $\varepsilon$  is the sum of variances of noise on all sensors, i.e.  $\varepsilon^2 = \sum_{l=1}^k \varepsilon_l^2$ . Combining this latter relationship with eq. (3.4) leads to determine the quasi-optimal tikhonov parameter found according to the discrepancy principle,  $\vartheta_{\text{dp}}^*$  solution of :

$$R(\vartheta_{\text{dp}}^*) = \sum_{i=1}^k \left( \frac{\vartheta_{\text{dp}}^* \check{\xi}_i}{\eta_i^2 + \vartheta_{\text{dp}}^*} \right)^2 - \varepsilon^2 = 0 \quad (3.6)$$

### 3.2.2 Numerical results

The numerical study shows that the solution of eq. (3.6) is almost independent of the dimension  $\Xi$  used in the parameterization of the flux. To do so, let the synthetic data be generated with the flux  $\bar{\varphi}(r_2) = \varphi_0 \left( \sin \left( \frac{\pi r_2}{2} \right) - 1 \right)$ ,  $\varphi_0 = -10^4$  W/m<sup>2</sup> and the thermal conductivity  $\lambda_c = 30$  W/m K. Five pointwise temperature measurements ( $k = 5$ ) are performed at locations  $r_1 = 0.9$ , and  $r_2 = 0.1, 0.3, 0.5, 0.7$  and  $0.9$ . The data is then perturbed according to a Gaussian white noise with variance  $\varepsilon_l^2$  ranging from  $10^{-4}$  to  $10^{-1}$  K<sup>2</sup>,  $\forall l = 1, \dots, k$ , yielding to  $\check{T}$ . The finite element method is used to solve the forward problem eq. (3.1) based on a regular mesh of the bounded domain  $\Omega$  and a discretization of the temperature  $T$  with Lagrange  $\mathcal{P}_1$  elements. The regular grid associated to  $\Omega$  is chosen sufficiently fine to ensure that errors due to the approximation method are negligible when compared to measurement errors. The basis functions  $(\Theta_l)_{l=1}^{\Xi}$  used in the parameterization of the unknown flux  $\varphi$  are also continuous first order Lagrange functions, i.e. linear functions per element satisfying  $\Theta_l(r_{2p}) = 1$  if and only if  $p = l$ .

Table 3.1 presents the quasi-optimal Tikhonov parameter  $\vartheta_{\text{dp}}^*$  for several uniform Lagrange parameterizations with  $\Xi$  ranging from 7 to 25, and several variances of noise in the data, from  $\varepsilon_l^2 = 10^{-4}$  to  $\varepsilon_l^2 = 10^{-1}$ . It is seen that, for a given level of noise, the Tikhonov parameter  $\vartheta_{\text{dp}}^*$  depends only very slightly on the flux discretization. The small fluctuations that remain may

come from all numerical approximations : finite element computations to build the state matrix  $A$ , the singular value decomposition  $W\Lambda V^t = A$  and the numerical optimization when solving the nonlinear problem eq. (3.6),  $\vartheta_{dp}^* = \arg \min R(\vartheta_{dp})$ . Moreover, it is worth noting the linear dependency of  $\vartheta_{dp}^*$  with the noise variance  $\varepsilon_l^2$  (Table 3.1).

Figure 3.1 presents the distance from the actual solution to the expected target distribution,  $E(\vartheta) = \frac{\|\varphi_\vartheta - \bar{\varphi}\|}{\|\bar{\varphi}\|}$ , as a function of the Tikhonov parameter  $\vartheta$ , for the variance noise  $\varepsilon_l = 10^{-3}$ ,  $\forall l$ . The norm was chosen to be defined as  $\|\varphi\| = \left( \int_{r_1=1} \varphi^2 dr_2 \right)^{\frac{1}{2}}$ . Results obtained for other noise variances gave similar curves to the one given in Figure 3.1 and are thus not presented here. Figure 3.1 shows that, for low Tikhonov parameters  $\vartheta$ , the distance from the solution to the target decreases when the control space dimension  $\Xi$  decreases. On the contrary, for large Tikhonov parameters, the distance from the solution to the target is independent of the control space dimension. In between, there is a plateau, and the minimum point is found to be independent of  $\Xi$ .

In order to illustrate the effect of regularization, different reconstructions are presented in Figure 3.2. This figure shows how parameterization influences the reconstructions (Favenne *et al.*, 2013; Balima *et al.*, 2013). It is also observed from Figure 3.2 that despite the very limited number of steady-state measurements (a study on the optimal number and location of inner pointwise measurements is out of the scope of this paper), the space-dependent heat flux could be retrieved for appropriate Tikhonov regularization combined with appropriate control space parameterization.

TAB. 3.1 Value of the Tikhonov parameter  $\vartheta_{dp}^*$  solution of (3.6) for different discretizations and different noise magnitudes.

$\Xi \setminus \varepsilon_l^2$	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$
<b>7</b>	$4.189 \times 10^{-11}$	$4.198 \times 10^{-10}$	$4.297 \times 10^{-9}$	$5.493 \times 10^{-8}$
<b>10</b>	$5.246 \times 10^{-11}$	$5.257 \times 10^{-10}$	$5.374 \times 10^{-9}$	$6.622 \times 10^{-8}$
<b>15</b>	$4.240 \times 10^{-11}$	$4.248 \times 10^{-10}$	$4.330 \times 10^{-9}$	$5.151 \times 10^{-8}$
<b>20</b>	$4.042 \times 10^{-11}$	$4.049 \times 10^{-10}$	$4.113 \times 10^{-9}$	$4.683 \times 10^{-8}$

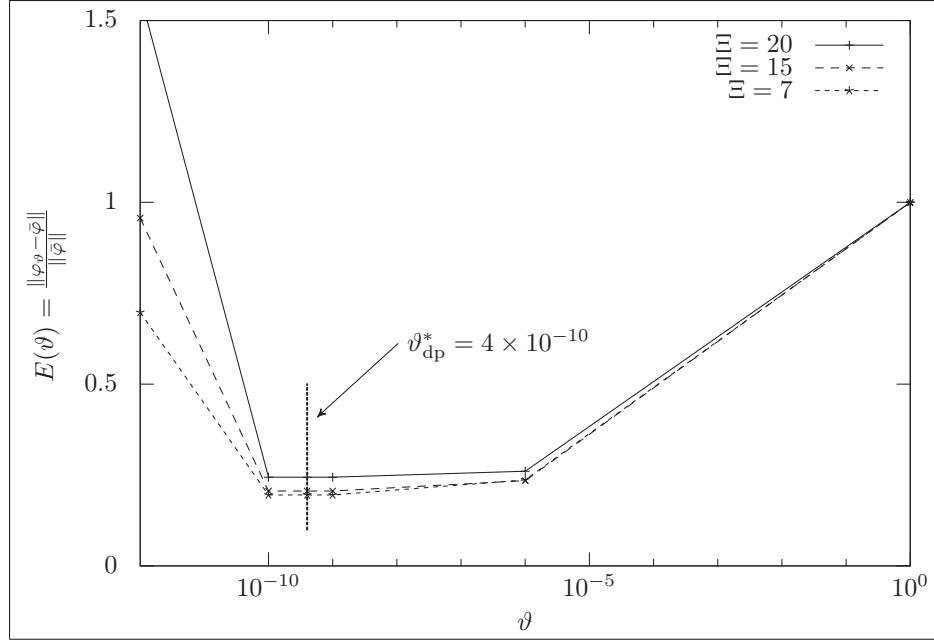


FIG. 3.1 Distance from the solution to the target  $E(\vartheta)$  as a function of the Tikhonov parameter  $\vartheta$ , for three control space dimensions  $\Xi$  equal to 7, 15 and 20. It is seen that  $\vartheta^* = \arg \min E(\vartheta)$  is independent of  $\Xi$ . The quasi-optimal Tikhonov parameter found according to the discrepancy principle has also been added.

### 3.3 Optical tomography

#### 3.3.1 Forward model and inversion setting

Solving for the DOT inverse problem is usually based on the minimization of a cost function which depends on the discrepancy between some measurements and the related predictions, the latter being a solution of the forward model which is, in the present case, based on the two-dimensional DA. The complex photon density,  $\varphi : \Omega \mapsto \mathbb{C}$ , is mathematically described by the DA model expressed in the frequency domain (Arridge, 1999) :

$$-\nabla \cdot ([2(\kappa + \sigma)]^{-1} \nabla \varphi) + (\kappa + \frac{2\pi i \omega}{c}) \varphi = 0 \text{ in } \Omega \quad (3.7)$$

$$\varphi + \frac{A}{2\pi^{-1}} [2(\kappa + \sigma)]^{-1} \nabla \varphi \cdot \mathbf{n} = \frac{f_\omega}{\pi^{-1}} \mathbb{1}_{[\zeta \in \partial\Omega_s]} \text{ on } \partial\Omega \quad (3.8)$$

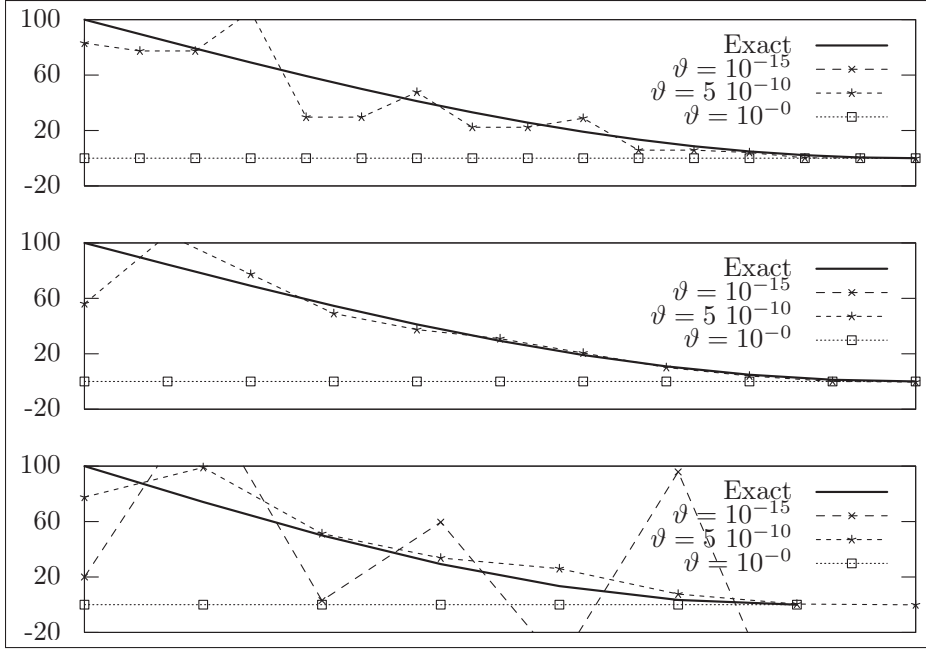


FIG. 3.2 Reconstructions  $\varphi_{\vartheta}(r_2)$  for  $\Xi = 15$  (top),  $\Xi = 10$  (middle) and  $\Xi = 7$  (bottom) for under-regularization ( $\vartheta = 10^{-15}$ ), over-regularization ( $\vartheta = 1$ ) and appropriate regularization ( $\vartheta = 5 \cdot 10^{-10} \approx \vartheta^*$ ). Note that points for  $\Xi = 15$  and  $\Xi = 10$  (over-parameterization) with  $\vartheta = 10^{-15}$  (under-regularization) are not presented because of divergence. The noise variance  $\varepsilon_l^2 = 0.001$  was used after generating the synthetic data.

where  $\kappa$  and  $\sigma$  are the absorption and reduced scattering coefficients, respectively,  $c$  is the speed of light in the medium,  $\mathbf{n}$  is the unit outward normal vector,  $f_{\omega}$  is the total inward flux of the amplitude modulated diffuse source at the frequency  $\omega$ ,  $\mathbb{1}_{[\cdot]}$  denotes the indicator function,  $\partial\Omega_s$  depicts the light source location,  $i$  is the imaginary unit and  $A$  is a parameter which characterizes the reflection at the boundary and can be derived from Fresnel's law if specular reflection is considered (Dehghani *et al.*, 2009a) or from experimental set-ups (Schweiger *et al.*, 1995).

The forward problem consists in solving eq. (3.7)-(3.8) for the photon density  $\varphi$  assuming the absorption coefficient  $\kappa$ , the reduced scattering coefficient  $\sigma$ , the location of the light source  $\partial\Omega_s$  and values of  $\omega$ ,  $c$ ,  $A$  and  $f_{\omega}$  are known. In contrast, the inverse problem consists in



estimating the space-dependent radiative properties  $\kappa$  and  $\sigma$  on the basis of photon density measurements  $\check{\phi}$  on  $\partial\Omega$ , under the assumption that the location of the light source  $\partial\Omega_s$  and values of  $\omega$ ,  $c$ ,  $A$  and  $f_\omega$  are known.

More specifically, the forward model leads to compute the state  $\phi$  that depends on radiative properties  $\gamma = (\kappa, \sigma)$ . The difference (in the least squares sense) between this density and the measured one is integrated to the cost function to be minimized  $j(\gamma) = \mathcal{J}(\phi)$  :

$$\mathcal{J}(\phi) = \frac{1}{2} \sum_{k=1}^K \sum_{d=1}^D \left| \frac{\phi_{(k;d)} - \check{\phi}_{(k;d)}}{\check{\phi}_{(k;d)}} \right|^2 \quad (3.9)$$

where the index  $k$  defines the source (test) number and the index  $d$  defines the detection number. The solution of the inverse problem requires an optimization problem formulation of the kind : “Find the functions  $\kappa^*(\mathbf{r})$  and  $\sigma^*(\mathbf{r})$  such that  $j(\kappa^*, \sigma^*) = \min j(\kappa, \sigma)$ ”.

### 3.3.2 Cost function derivative

In order to derive the GN algorithm, one first needs to recall basic definitions of directional derivatives. Following (Lions et Faurre, 1982), let us denote the point  $(\kappa, \sigma) = \gamma \in \Lambda^2 \subset [L^2(\Omega)]^2$  and directions  $\eta$  and  $\zeta \in \Lambda$ . The directional derivative of the state  $\phi$  at point  $\gamma$  towards  $\eta$  is

$$\phi'(\gamma; \eta) := \lim_{\varepsilon \rightarrow 0^+} \frac{\phi(\gamma + \varepsilon \eta) - \phi(\gamma)}{\varepsilon} \quad (3.10)$$

An analogous definition can be stated for the directional derivative of the cost function along with its second-order directional derivative such that :

$$j''(\gamma; \eta, \zeta) := \lim_{\varepsilon \rightarrow 0^+} \frac{j'(\gamma + \varepsilon \zeta; \eta) - j'(\gamma; \eta)}{\varepsilon} \quad (3.11)$$

Next, the operators involved in Newton's method are extracted through the following equations :

$$j'(\gamma; \eta) = (\nabla j(\gamma), \eta)_{L^2(\Omega)} \quad (3.12)$$

$$j''(\gamma; \eta, \zeta) = (\nabla^2 j(\gamma) \eta, \zeta)_{L^2(\Omega)} \quad (3.13)$$

with the inner product is defined as  $(\gamma, \eta)_{L^2(\Omega)} = \int_{\Omega} \gamma \eta \, d\mathbf{r}$ .

### 3.3.3 Parameterization

The solution of the optimization problem relies on iterative minimization algorithms for which control and state spaces must be finite in practice. To do so, let the functions to be determined  $\lambda = \kappa, \sigma$  and  $\varphi$  be approximated using finite element basis functions  $(\psi_{\xi})_{\xi=1}^{\Xi_{\lambda}}$  and  $(\Theta_{\xi})_{\xi=1}^{\Xi_{\varphi}}$ , respectively :

$$\lambda(\mathbf{r}) = \lambda_{\text{ap}} \sum_{\xi=1}^{\Xi_{\lambda}} \psi_{\xi}(\mathbf{r}) \tilde{\lambda}(\mathbf{r}_{\xi}), \quad \lambda = \kappa, \sigma \quad (3.14)$$

$$\varphi(\mathbf{r}) = \sum_{\xi=1}^{\Xi_{\varphi}} \Theta_{\xi}(\mathbf{r}) \varphi(\mathbf{r}_{\xi}) \quad (3.15)$$

where it should be noted that the dimension of the control space  $\Xi_{\lambda}$  is chosen less or equal to that of the state space  $\Xi_{\varphi}$ . As a consequence, projections of functions of the expression (3.14) into the state functional space must be used to make the solution of the system (3.7)-(3.8) possible. Moreover, parameters are adimensionalized with *a priori* properties  $\lambda_{\text{ap}} = \kappa_{\text{ap}}, \sigma_{\text{ap}}$  such that  $\lambda = \lambda_{\text{ap}} \tilde{\lambda}$ , in order to make both functions  $\tilde{\lambda} = \varkappa(\mathbf{r}), \varsigma(\mathbf{r})$  to be searched about unity.

Let us note  $\tilde{\gamma} = (\varkappa, \varsigma)$ . Parameterization of radiative property map functions allows the construction of the GN matrix system which is written as :

$$\tilde{\nabla}^2 j(\tilde{\gamma}) \delta \tilde{\gamma} = -\tilde{\nabla} j(\tilde{\gamma}) \quad (3.16)$$

with state second-order derivatives assumed to be negligible when compared to first-order state derivatives and where vector  $\tilde{\nabla}j$  and matrix  $\tilde{\nabla}^2j$  represent continuous gradient  $\nabla j$  and approached Hessian matrix  $\nabla^2j$  of (3.12)-(3.13) decomposed on finite element basis functions  $(\psi_\xi)_{\xi=1}^{\Xi_\lambda}$ . Finally, the GN matrix system to be solved in order to update radiative properties at each iteration is expressed as :

$$\text{Re} \left( \mathcal{S}^\top \overline{\mathcal{S}} \right) \delta \tilde{\gamma} = -\text{Re} \left( \mathcal{S}^\top \overline{\mathcal{R}} \right) \quad (3.17)$$

where  $\mathcal{S}^\top$  and  $\overline{\mathcal{S}}$  are the transposed and the conjugate of  $\mathcal{S}$ , respectively, and  $\text{Re}(\cdot)$  denotes the real part of the imaginary vector or matrix. The forward model behaving nonlinearly with respect to properties  $\tilde{\gamma} = (\varkappa, \varsigma)$ , the increment  $\delta \tilde{\gamma}^l = \tilde{\gamma}^l - \tilde{\gamma}^{l-1}$  given by the solution of the GN equation (3.17) is solved several times until convergence is attained. With the nomenclature defined earlier, the expressions for  $\mathcal{S}$  and  $\mathcal{R}$  are obtained such that :

$$\mathcal{S} = \left( \frac{(\varphi'_\varkappa)_{(k;d)}^\xi}{|\check{\Phi}_{(k;d)}|}, \frac{(\varphi'_\varsigma)_{(k;d)}^\xi}{|\check{\Phi}_{(k;d)}|} \right)_{\substack{\xi=1,\dots,\Xi_\lambda \\ (k;d)=(1,\dots,K;1,\dots,D)}} \quad (3.18)$$

$$\mathcal{R} = \left( \frac{\varphi_{(k;d)} - \check{\Phi}_{(k;d)}}{|\check{\Phi}_{(k;d)}|} \right)_{(k;d)=(1,\dots,K;1,\dots,D)} \quad (3.19)$$

where  $K$  and  $D$  are the number of sources and sensors, respectively.  $\mathcal{S} \in \mathbb{C}^{(K \times D) \times 2\Xi_\lambda}$ ,  $\mathcal{R} \in \mathbb{C}^{(K \times D)}$ , and derivative functions  $\varphi'_\varkappa$  and  $\varphi'_\varsigma$  can be obtained directly by differentiating the forward model (3.7)-(3.8) with respect to the adimensionalized radiative properties  $\varkappa$  and  $\varsigma$ , respectively.  $\varphi'_\varkappa$  and  $\varphi'_\varsigma$  are solution of the following partial differential equations :

$$-\nabla \cdot ([2(\kappa + \sigma)]^{-1} \nabla \varphi'_\varkappa) + (\kappa + \frac{2\pi i \omega}{c}) \varphi'_\varkappa = -\kappa_{\text{ap}} \varkappa' \varphi - \kappa_{\text{ap}} \nabla \cdot \left( \frac{\varkappa' \nabla \varphi}{2(\kappa + \sigma)^2} \right) \text{ in } \Omega \quad (3.20)$$

$$\varphi'_\varkappa + \frac{A}{2\pi^{-1}} [2(\kappa + \sigma)]^{-1} \nabla \varphi'_\varkappa \cdot \mathbf{n} = \kappa_{\text{ap}} \frac{A}{2\pi^{-1}} \frac{\varkappa' \nabla \varphi}{2(\kappa + \sigma)^2} \nabla \varphi \cdot \mathbf{n} \text{ on } \partial\Omega \quad (3.21)$$

and

$$-\nabla \cdot ([2(\kappa + \sigma)]^{-1} \nabla \phi'_\zeta) + (\kappa + \frac{2\pi i \omega}{c}) \phi'_\zeta = -\sigma_{\text{ap}} \nabla \cdot \left( \frac{\zeta' \nabla \phi}{2(\kappa + \sigma)^2} \right) \text{ in } \Omega \quad (3.22)$$

$$\phi'_\zeta + \frac{A}{2\pi^{-1}} [2(\kappa + \sigma)]^{-1} \nabla \phi'_\zeta \cdot \mathbf{n} = \sigma_{\text{ap}} \frac{A}{2\pi^{-1}} \frac{\zeta' \nabla \phi}{2(\kappa + \sigma)^2} \nabla \phi \cdot \mathbf{n} \text{ on } \partial\Omega \quad (3.23)$$

where  $\varkappa'$ ,  $\zeta'$  denote the perturbations. Algorithm 3.1 gives some explanations about the construction of vector  $\mathcal{R}$  and matrix  $\mathcal{S}$  involved in (3.17).

Algorithme 3.1 : Construction of vector  $\mathcal{R}$  and matrix  $\mathcal{S}$  involved in (3.17).

**Input :**  $\tilde{\gamma} = (\varkappa, \zeta)$ ;

Compute  $\kappa = \kappa_{\text{ap}} \varkappa$ ,  $\sigma = \sigma_{\text{ap}} \zeta$ ;

**for**  $k \leftarrow 1$  **to**  $K$  **do**

    Compute the state variable  $\phi$  solving (3.7)-(3.8);

    Extract the complex values of the photon density  $\phi$  at the nodes of the sensors and fill the corresponding part to the source  $k$  in vector  $\mathcal{R}$  (3.19);

**for**  $\xi \leftarrow 1$  **to**  $\Xi_\lambda$  **do**

        Set perturbations  $\varkappa'$  and  $\zeta'$  to the finite element basis function  $\psi_\xi$ ;

        Compute  $\phi'_\varkappa$  and  $\phi'_\zeta$  solving (3.20)-(3.21) and (3.22)-(3.23), respectively;

        Extract the complex values of  $\phi'_\varkappa$  and  $\phi'_\zeta$  at the nodes of the sensors and fill the corresponding part to the source  $k$  and the perturbation  $\xi$  in matrix  $\mathcal{S}$  (3.18);

**end**

**end**

**return**  $\mathcal{R}, \mathcal{S}$ ;

### 3.3.4 Tikhonov regularization

The Tikhonov regularization method consists in adding a penalization term  $\vartheta \mathcal{J}^+$  which is based on priors to the cost to be minimized (3.9) :

$$\mathcal{J}^+(\varkappa, \zeta) = \frac{1}{2} \sum_{\tilde{\lambda} = \varkappa, \zeta} \int_{\Omega} \left( \tilde{\lambda} - \tilde{\lambda}_{\text{ap}} \right)^2 \text{d}\mathbf{r} \quad (3.24)$$

$\vartheta$  being the Tikhonov parameter, and  $\tilde{\lambda}_{\text{ap}}$  being equal to 1 as explained above. This has been successfully used in the DOT area e.g. by (Paulsen et Jiang, 1995; Schweiger *et al.*, 2005; Niu *et al.*, 2008; Tarvainen *et al.*, 2008a; Dehghani *et al.*, 2009a,b) when using the GN optimizer. Doing so, the optimization problem becomes “Find the functions  $\kappa_{\vartheta}^*(\mathbf{r})$  and  $\varsigma_{\vartheta}^*(\mathbf{r})$  such that  $j_{\vartheta}(\kappa_{\vartheta}^*, \varsigma_{\vartheta}^*) = \min j_{\vartheta}(\kappa, \varsigma)$ ” with  $j_{\vartheta} = \mathcal{J} + \vartheta \mathcal{J}^+$ .

The simultaneous use of several regularizations together has been studied in (Favenne *et al.*, 2013; Balima *et al.*, 2013), among which the reduction of the control space dimension (Chavent, 2010) with Tikhonov regularization. It is shown here that the control space reduction modifies only slightly the optimal Tikhonov parameter.

The control space mesh is assumed to be chosen sufficiently fine to ensure that the spatial fluctuations of the radiative properties are well taken into account. Thus, the projections from the control functional space to the state functional space are such that the state is almost independent of the number of degrees of freedom  $\Xi_{\lambda}$ , thus the cost function  $\mathcal{J}$  is also almost independent of  $\Xi_{\lambda}$ . Next, to simplify the presentation, let us consider only one parameter, say  $\kappa$ .

**Proposition 1.** *Under the following assumptions : H1)  $\psi_{\xi}$  are basis functions constant per element on a regular mesh, i.e. Lagrange  $\mathcal{P}_0$  elements are used for the discretization of the function  $\kappa$ , H2)  $\kappa \sim \mathcal{N}(1, \sigma_{\kappa}^2)$ , then the optimal Tikhonov parameter is asymptotically independent of the dimension  $\Xi_{\lambda}$ .*

**Proof.** Let  $\kappa_{\vartheta}^*$  be a solution of the optimization problem

$$\kappa_{\vartheta}^* = \arg \min_{\kappa} (\mathcal{J}(\kappa) + \vartheta \mathcal{J}^+(\kappa)) \quad (3.25)$$

with  $\vartheta$  representing the best compromise between both costs  $\mathcal{J}(\kappa)$  and  $\mathcal{J}^+(\kappa)$ . Then, showing  $\mathcal{J}^+$  does not depend on the parameterization yields  $\vartheta$  independent of  $\Xi_{\lambda}$ . Let  $\tilde{\vartheta} = \sigma_{\kappa}^2 \vartheta$  and  $\Delta x = |\Omega|/\Xi_{\lambda}$  be the area of one element of the regular mesh. With H1), we have  $\vartheta \mathcal{J}^+(\kappa) = \frac{\tilde{\vartheta}|\Omega|}{2} \frac{1}{\Xi_{\lambda}} \sum_{i=1}^{\Xi_{\lambda}} \left( \frac{\kappa_i - 1}{\sigma_{\kappa}} \right)^2$ . With H2),  $\frac{\kappa_i - 1}{\sigma_{\kappa}} \sim \mathcal{N}(0, 1) \forall i$ , thus  $\Upsilon = \sum_{i=1}^{\Xi_{\lambda}} \left( \frac{\kappa_i - 1}{\sigma_{\kappa}} \right)^2 \sim$

$\chi_{\Xi_\lambda}^2$ . Then, using the fact that  $\mathbf{E}(\Upsilon) = \Xi_\lambda$ , we have  $\mathbf{E}(\mathcal{J}^+) = \frac{\sigma_\kappa^2 |\Omega|}{2}$  which is independent of  $\Xi_\lambda$ . This completes the proof.

**Remark 11.** This demonstration can be easily extended to several parameters, say with a function of the type of (3.24) assuming that  $\sigma_\zeta \approx \sigma_\kappa$ .

**Remark 12.** It could also be extended to other finite element parameterizations, but with much heavier calculations.

### 3.3.5 Numerical validation

The ability to determine on a coarse mesh a reasonable parameter for the Tikhonov regularization is illustrated on a circular test medium. The rationale behind this choice is to reduce computational time required for getting a Tikhonov parameter so that the solution is close enough to the true solution, at least from an engineering point of view.

Synthetic data and predicted values are collected with the help of 8 equally distributed sources and sensors about the disk perimeter ( $K = 8$ ). Sources and sensors are alternatively distributed as schematically depicted in Figure 3.3. Each sensor contains 7 pointwise photon density measurements ( $D = 56$ ). Radiative property maps to be recovered are also given in Figure 3.3. The areas to be detected are located at about 2.5 cm in opposite locations at  $45^\circ$  and  $225^\circ$ , respectively. The defects are  $4/3$  cm in diameters while the whole medium has a 4 cm radius. Each probe is 0.8 cm in length. The other physical parameters involved in (3.7)-(3.8) are fixed to :  $n = 1.4$ ,  $\omega = 1 \times 10^8 \text{ s}^{-1}$ ,  $c = c_0/n$  with  $c_0 = 3 \times 10^{10} \text{ cm.s}^{-1}$ ,  $f_\omega = 0.1 \text{ W.cm}^{-2}$  and  $A$  is derived from Fresnel's law (see (Dehghani *et al.*, 2009a)).

The radiative parameters are initialized to the radiative properties of the background considered as priors :  $\kappa^0 = \kappa_{\text{ap}} = 0.8 \text{ cm}^{-1}$  and  $\sigma^0 = \sigma_{\text{ap}} = 20 \text{ cm}^{-1}$  all over the medium  $\Omega$ . Thus, initial control variables  $\kappa^0 = \kappa^0 / \kappa_{\text{ap}}$  and  $\zeta^0 = \sigma^0 / \sigma_{\text{ap}}$  are initialized to 1. Synthetic data  $\check{\phi}$  and initial adimensionalized radiative properties  $\kappa$  and  $\zeta$  are inputs of the inverse problem. The way in which synthetic data are generated is given hereafter.

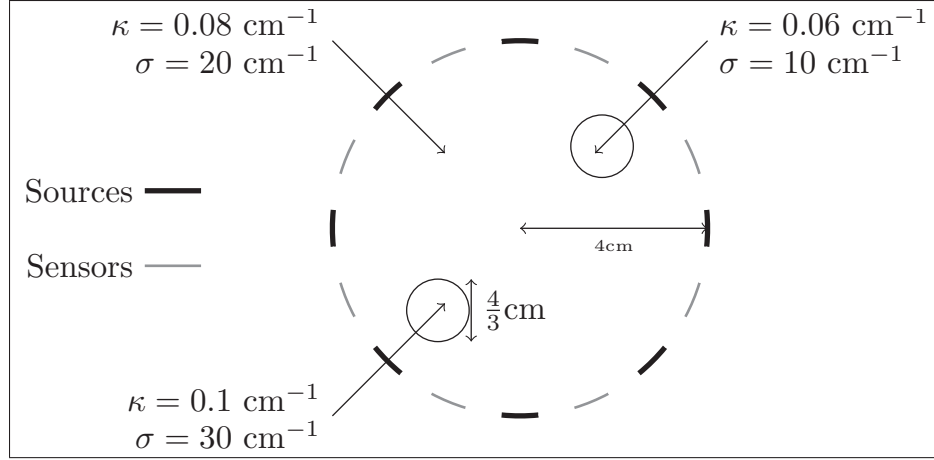


FIG. 3.3 Test medium geometry representation. Eight sources are located on the boundary. For each source (which constitutes a test), the emerging radiation is measured on all sensors.

Within the inverse procedure, the forward model (3.7)-(3.8) is solved by a finite element method with Lagrange  $\mathcal{P}_1$  elements, that is with  $\varphi \in \mathcal{V}_h^\varphi = \mathcal{V}(\mathcal{M}_h^\varphi, \mathcal{P}_1)$ . The state mesh  $\mathcal{M}_h^\varphi$  is chosen fine enough to ensure that numerical predictions can fit the synthetic data. Synthetic data  $\check{\varphi}$  are also generated with the finite element method with Lagrange  $\mathcal{P}_1$  elements but using a much finer mesh  $\mathcal{M}_h^{\check{\varphi}}$  in order to avoid the inverse crime (Kaipio et Somersalo, 2005). After projecting the synthetic data in the functional space  $\mathcal{V}_h^\varphi$ ,  $\check{\varphi}$  are corrupted with a multiplicative white Gaussian noise of signal-to-noise ratio  $\text{SNR} = 30$  dB :  $\check{\varphi}_{\text{noisy}} = \check{\varphi} \left( 1 + 10^{-\text{SNR}/10} \times \varepsilon \right)$  where  $\varepsilon \sim \mathcal{N}(0, 1)$ .

Next, the control space  $\mathcal{V}_h^\lambda = \mathcal{V}(\mathcal{M}_h^\lambda, \mathcal{P}_1)$  for  $\lambda = \kappa, \sigma$  relies on a more or less coarse mesh  $\mathcal{M}_h^\lambda$  according to cases (see Table 3.2). The FreeFem++ environment has been used to perform all computations (Hecht, 2012).

TAB. 3.2 Dimensions of finite element spaces for  $\varphi$ ,  $\check{\varphi}$  and  $\lambda$ .

	$\Xi_\lambda = \dim \mathcal{V}_h^\lambda$	$\dim \mathcal{V}_h^\varphi$	$\dim \mathcal{V}_h^{\check{\varphi}}$
<b>case 1</b>	2924	2924	5336
<b>case 2</b>	1382		
<b>case 3</b>	384		

The error curves  $E(\vartheta)$  are built at the 20<sup>th</sup> iteration for the three cases described in Table 3.2 : a fine mesh for the state ( $\dim \mathcal{V}_h^\vartheta = 2924$ ), and successively a fine mesh, a medium mesh and a coarse mesh with respectively  $\dim \mathcal{V}_h^\lambda$  equal to 2924, 1382 and 384 nodes for the parameters. The three error curves are presented all together in Figure 3.4. Figure 3.4 first shows that for  $\vartheta > 0.4$ , the calculated errors are independent of the mesh coarseness. Otherwise, for  $\vartheta < 0.3$ , it is seen that the reduction of the control space reduces dramatically the errors. Such reduction of control space thus regularizes the inverse problem preconditioning the optimization (Chavent, 2010). In between, Figure 3.4 shows that the optimal Tikhonov parameters are found to be quasi-independent of the number of degrees of freedom of the finite control space. Within the range  $[10^{-2}, 10^{+2}]$ , the optimal Tikhonov parameter  $\vartheta^*$  is found to be equal to 0.3 for  $\Xi_\lambda = 384$  and 0.4 for  $\Xi_\lambda = 1382$  and 2924. This means that the minimal residual is indeed quasi-independent of the mesh coarsening, i.e. of  $\dim \mathcal{V}_h^\lambda$ . This is what has been shown in Proposition 1.

Table 3.3 also provides the relative CPU time needed to build such curves for cases 1, 2 and 3 described in Table 3.2 : the construction of this error curve with the coarsest mesh needs, in this case, roughly the seventeenth of the initial CPU time related to the construction with the finest mesh. This can be easily understood by examining the construction scheme of the GN matrix system presented above (see Algorithm 3.1).

Figures 3.5 and 3.6 present the reconstructions of both  $\kappa$  and  $\sigma$  obtained with the GN optimizer at the 20<sup>th</sup> iteration, with  $\Xi_\lambda = 2924$ , and with the Tikhonov weight equal to 0.2 and 0.4, respectively. Figure 3.5 shows the divergence of the reconstruction with negative values for  $\sigma$  due to under-regularization : the weight parameter  $\vartheta$  is too low. Figure 3.6 presents the same reconstruction but with the Tikhonov parameter  $\vartheta$  chosen at the minimum of the error curve. Reconstructions are good even in the presence of a 30 dB noise.

Two supplementary numerical tests are considered in order to investigate the behaviour of the inverse method with respect to the initial guess. As previously, initial control variables  $\varkappa^0$  and  $\varsigma^0$  are initialized to 1 so that the initial radiative properties are equal to the *a priori* properties



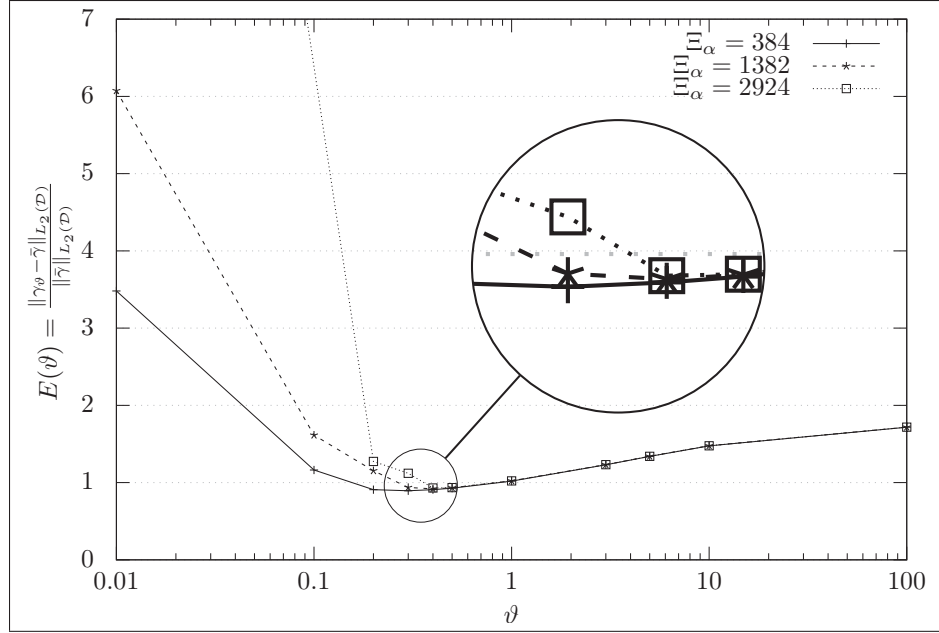


FIG. 3.4 Distance from the solution to the target  $E(\vartheta)$  as a function of the Tikhonov parameter  $\vartheta$ , for  $(\kappa_{\text{ap}}, \sigma_{\text{ap}}) = (0.08, 20) \text{ cm}^{-1}$  and three control space dimensions  $\Xi_\lambda$  equal to 384, 1382 and 2924. It is seen that  $\vartheta^* = \arg \min E(\vartheta)$  is quasi-independent of  $\Xi$ .

TAB. 3.3 CPU time comparisons for the generation of the error curves  $E(\vartheta)$ .

	Case 1	Case 2	Case 3
CPU time	1	0.38	0.09

$\kappa_{\text{ap}}$  and  $\sigma_{\text{ap}}$ . For the first test, *a priori* properties are fixed to  $(\kappa_{\text{ap}}, \sigma_{\text{ap}}) = (0.088, 24) \text{ cm}^{-1}$  (ten and twenty percent larger than radiative properties of the background medium for the absorption and reduced scattering coefficients, respectively). The error curves show that the optimal Tikhonov parameter  $\vartheta^*$  is found to be equal to 1.1, 1.3 and 1.4 for  $\Xi_\lambda = 384, 1382$  and 2924, respectively. This illustrates again the quasi-independence of the optimal Tikhonov parameter with the control space parameterization. However, it should be noted that, as for the previous case where *a priori* properties equal the radiative properties of the background, the optimal Tikhonov parameter is found to be slightly smaller for  $\Xi_\lambda = 384$  (1382) than for  $\Xi_\lambda = 1382$  (2924). This can be explained by the fact that the reduction of the control space dimension

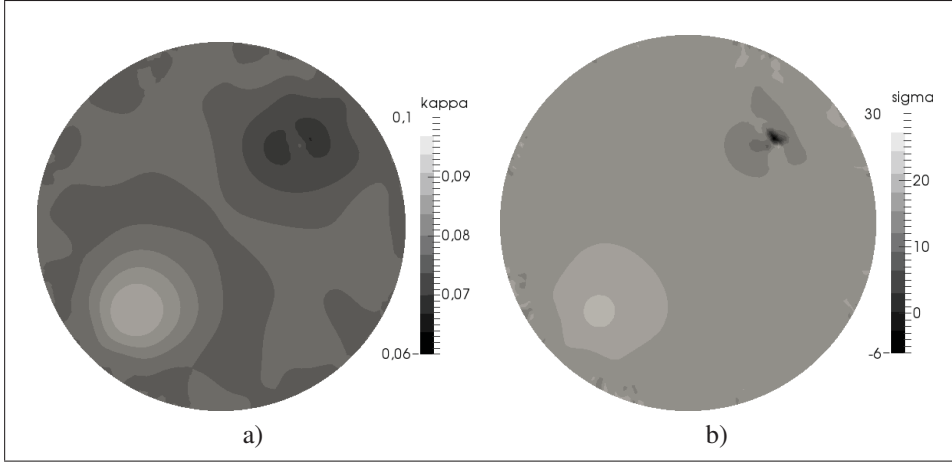


FIG. 3.5 Reconstruction of  $\kappa$  (left) and  $\sigma$  (right) with  $\vartheta = 0.2$ ,  $(\kappa_{\text{ap}}, \sigma_{\text{ap}}) = (0.08, 20) \text{ cm}^{-1}$  and  $\Xi_\lambda = 2924$ . It should be noted the divergence for the reduced diffusion coefficient.

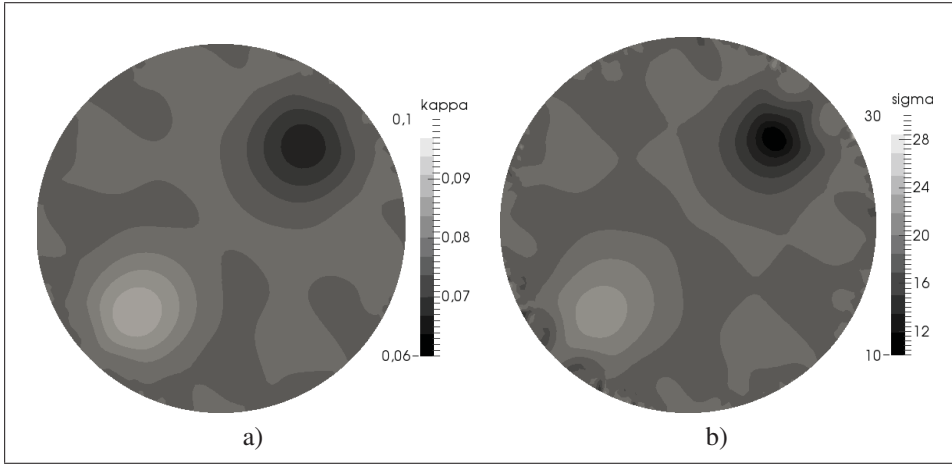


FIG. 3.6 Reconstruction of  $\kappa$  (left) and  $\sigma$  (right) with  $\vartheta = 0.4$ ,  $(\kappa_{\text{ap}}, \sigma_{\text{ap}}) = (0.08, 20) \text{ cm}^{-1}$  and  $\Xi_\lambda = 2924$ .

acts as a regularization method so that the smaller the control space dimension  $\Xi_\lambda$ , the smaller the optimal Tikhonov parameter  $\vartheta^*$ . As a result, the optimal Tikhonov parameter provided by a strongly reduced control space should not be considered as such for the original control space, but has to be interpreted as a very good indicator : the optimal Tikhonov parameter for the original control space will be slightly larger. Figure 3.7 presents the reconstructions of both  $\kappa$  and  $\sigma$  obtained with the GN optimizer at the 20<sup>th</sup> iteration, with  $\vartheta^* = 1.4$  and  $\Xi_\lambda = 2924$ . It is

observed that the reconstructions are less accurate than those obtained when *a priori* properties equal the radiative properties of the background, but the inclusions can still be distinguished.

Finally, for the second test, the *a priori* properties are fixed to  $(\kappa_{\text{ap}}, \sigma_{\text{ap}}) = (0.096, 28) \text{ cm}^{-1}$  (twenty and forty percent larger than radiative properties of the background medium for the absorption and reduced scattering coefficients, respectively). It was observed that the error curves associated to this initial guess are very irregular so that it is difficult to find an optimal Tikhonov parameter. In fact, the results show that the GN optimizer fails to reconstruct the target properties whatever the value of the Tikhonov parameter. The initial guess is too far from the target so that the descent direction of the GN algorithm is never pointed towards the minimum and leads inevitably to very inaccurate reconstructions whatever the Tikhonov weight. Thus, for such choice of the initial guess, the use of more sophisticated methods, such as the globally convergent ones developed in (Su *et al.*, 2006; Shan *et al.*, 2008a,b; Beilina et Klivanov, 2008, 2012; Su *et al.*, 2014), become necessary.

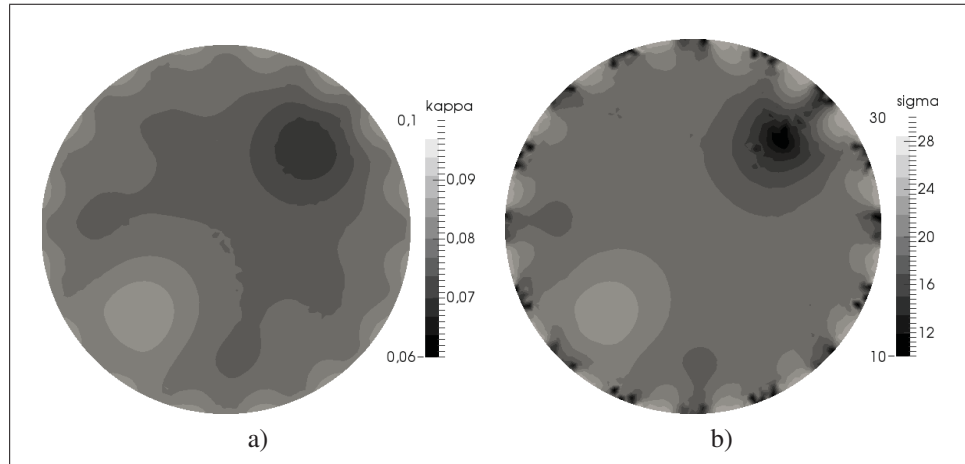


FIG. 3.7 Reconstruction of  $\kappa$  (left) and  $\sigma$  (right) with  $\vartheta = 1.4$ ,  $(\kappa_{\text{ap}}, \sigma_{\text{ap}}) = (0.088, 24) \text{ cm}^{-1}$  and  $\Xi_{\lambda} = 2924$ .

### 3.4 Conclusion

This paper has demonstrated that the optimal Tikhonov parameter, defined as that which minimizes the distance between the actual solution and the noise-free solution, is almost independent of the control space parameterization. This has first been verified on a standard linear inverse heat conduction problem in which, in phase one, the Tikhonov parameter has been approximated with the singular value decomposition and the maximum discrepancy principle and then, in a second phase, in the sense of the optimal Tikhonov parameter previously defined. Next, this validation has been extended to the nonlinear inverse problem of diffuse optical tomography. The error curves with respect to the Tikhonov parameter showed that their minimum are quasi-independent of the control space parameterization. This may result in CPU time reduction when searching for the optimal Tikhonov parameter, whatever the method used to find it. Of course, this is highly recommended for large size (three-dimensional) objects when the GN optimizer is chosen. Finally, note that the two inverse problems dealt with in this paper were under-determined, as it is the case in most space-dependent function estimation problems. Consequently, the theory has been numerically validated only on these two specific inverse problems. This may be viewed as a limitation, but the need of control space dimension reduction is much less useful for over-determined inverse problems than for space-dependent ones which are usually under-determined.

## CHAPITRE 4

### ARTICLE 3 : A WAVELET MULTI-SCALE METHOD FOR THE INVERSE PROBLEM OF DIFFUSE OPTICAL TOMOGRAPHY

Fabien Dubot<sup>1,2</sup>, Yann Favennec<sup>2</sup>, Benoit Rousseau<sup>2</sup>, Daniel Rousse<sup>1</sup>

<sup>1</sup> Chaire de recherche industrielle t3e, École de technologie supérieure,

1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

<sup>2</sup> Laboratoire de thermocinétique de Nantes, UMR CNRS 6607,

Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France

Article accepté dans la revue « Journal of Computational and Applied Mathematics » en janvier 2015.

#### Abstract

This paper deals with the estimation of radiative property distributions of participating media from a set of light sources and sensors located on the boundaries of the medium. This is the so-called diffuse optical tomography problem. Such a non-linear ill-posed inverse problem is solved through the minimization of a cost function which depends on the discrepancy, in a least-square sense, between some measurements and associated predictions. In the present case, predictions are based on the diffuse approximation model in the frequency domain while the optimization problem is solved by the L-BFGS algorithm. To cope with the local convergence property of the optimizer and the presence of numerous local minima in the cost function, a wavelet multi-scale method associated with the L-BFGS method is developed, implemented, and validated. This method relies on a reformulation of the original inverse problem into a sequence of sub-inverse problems of different scales using wavelet transform, from the largest scale to the smallest one. Numerical results show that the proposed method brings more stability with respect to the ordinary L-BFGS method and enhances the reconstructed images for most of initial guesses of radiative properties.

## 4.1 Introduction

The diffuse optical tomography (DOT) problem is an inverse problem in which the spatial distributions of absorption and scattering coefficients of a participating medium are sought from a set of light sources and sensors located on the frontier of the probed medium (Arridge, 1999). The main application of the DOT is imaging biological tissues based on the fact that knowledge of the radiative properties provides information on the physiopathological condition of these tissues. A thorough review of work on applications of the DOT in the biomedical domain can be found in (Gibson *et al.*, 2005).

Such a non-linear ill-posed inverse problem is solved through the minimization of a cost function which depends on the discrepancy, in a least-square sense, between some measurements and associated predictions. The methods employed to solve this problem include the non-linear conjugate-gradient method (Arridge et Schweiger, 1998), Gauss-Newton (GN) based methods (Schweiger *et al.*, 2005; Niu *et al.*, 2008; Dehghani *et al.*, 2009b; Tarvainen *et al.*, 2011), the L-BFGS method (Klose et Hielscher, 2003; Ren *et al.*, 2006; Balima *et al.*, 2011a, 2013; Favennec *et al.*, 2013), shape-based reconstruction method (Zacharopoulos *et al.*, 2006; Arridge *et al.*, 2008) or, in a Bayesian framework, the approximation error method (Arridge *et al.*, 2006; Kaipio et Somersalo, 2005). Regarding the first three listed methods, some problems remain to be overcome such as the stability with respect to the initial guesses and the blurring effect of the reconstructed images due to the need for relatively strong regularization tools (Cao *et al.*, 2007; Arridge *et al.*, 2008).

A relatively new method has emerged in the field of inversion, namely the wavelet multi-scale method (Liu, 1993; Ding *et al.*, 2009; Zhang *et al.*, 2006; He et Han, 2009; Fu *et al.*, 2007; Lei *et al.*, 2011; Fu *et al.*, 2013; Zhao *et al.*, 2014). This method relies on a reformulation of the original inverse problem into a sequence of sub-inverse problems of different scales using wavelet transform, from the largest scale to the smallest one. Successful applications of this method include the inversion of the Maxwell equations (Ding *et al.*, 2009), the identification of space-dependent porosity in fluid-saturated porous media (Zhang *et al.*, 2006; He et Han, 2009), the

inversion of the two-dimensional acoustic wave equation (Fu *et al.*, 2007), the reconstruction of permittivity distribution by the inversion of the electrical capacitance tomography model (Lei *et al.*, 2011), the parameter estimation of elliptical partial differential equations (Liu, 1993; Fu *et al.*, 2013) and the identification of space-dependent permeability in a nonlinear diffusion equation (Zhao *et al.*, 2014). It is shown in these papers that the wavelet multi-scale method can enhance stability of inversion, accelerate convergence and cope with the presence of local minima to reach the global minimum. The rationale behind this success is that the cost function shows stronger convexity and has less minima at larger scale such that the global minimum can be achieved. Best results can thus be obtained, using this solution to the initialization of the optimization problem, at the lower scale until finding the minimum at the original scale.

This paper is dedicated to the application of a wavelet multi-scale method associated with the L-BFGS method to the specific inverse problem of DOT and aims to improve the efficiency of the ordinary L-BFGS algorithm. To the best knowledge of the authors, this is the first time that the L-BFGS algorithm is coupled with the multi-scale method in solving the DOT inverse problem.

The paper is organized as follows : section 4.2 shortly summarizes the models used in the area of optical tomography (OT), namely the radiative transfer equation (RTE) and the related diffuse approximation (DA) model ; section 4.3 introduces the cost function, provides the expression of the continuous cost function gradient and defines the optimization problem ; section 4.4 describes the L-BFGS algorithm ; section 4.5 presents an introduction to the wavelet theory and the proposed wavelet multi-scale method ; and section 4.6 provides selected numerical results to assess the proposed method.

But before presenting the work carried out in the paper, some clarification is mandatory to fully understand the originality of the work. Indeed, although some papers found in the literature deal with the DOT problem using the wavelet tool such as (Nath *et al.*, 1999; Zhu *et al.*, 1997a,b, 1998), these works show significant differences with the current research :

- In (Nath *et al.*, 1999), the wavelet transform is used to denoise and compress noisy experimental data before performing the reconstruction. In this paper, the use of the wavelet transform is not limited to the data filtering, the wavelet transform is used in all stages of the inverse problem.
- In (Zhu *et al.*, 1997a,b, 1998), the DOT problem is solved through the solution of a linear perturbation equation in which the main ingredient is a matrix of weights that are essentially the derivatives of the detector readings with respect to the optical coefficients in the reference medium. Several iterative algorithms have been used to solve this matrix system, including the conjugate gradient descent (Zhu *et al.*, 1997a) and the total least squares (Zhu *et al.*, 1997b, 1998). In all these algorithms, the multi-scale approach consists in multiplying both terms of the matrix system by wavelet transform matrices at each scale to obtain solutions at the different scales. In the present paper, the methodology is totally different : the DOT problem is solved through the minimization of a least squares cost function by the L-BFGS algorithm for which the main ingredient is the cost function gradient. Thus, the multi-scale approach is mainly based on the representation of this gradient at the different scales.
- In (Zhu *et al.*, 1997a,b, 1998), the steady-state DA was considered while the DA in the frequency domain is used in the present case. As a result, only the absorption coefficient was reconstructed in (Zhu *et al.*, 1997a,b, 1998) assuming the reduced scattering coefficient known within the media, while the simultaneous reconstruction of both the absorption and reduced scattering coefficients is performed in this paper, which represents an additional difficulty.

As a final comment about the current paper, it is worth mentioning that a wavelet multi-scale method associated with the L-BFGS algorithm is compared to the ordinary L-BFGS method. Comparisons with other methods found in the literature such as the GN method is out of the scope of the paper. However, let us point-out that the L-BFGS method has proven its worth in the field of the OT as shown by the following papers (Klose et Hielscher, 2003; Ren *et al.*, 2006; Balima *et al.*, 2011a, 2013).



## 4.2 The diffuse optical tomography problem

In the framework of radiative heat transfer in participating medium, the quantity of interest is the radiative intensity  $I$ . The equation describing the spectral radiative intensity distribution in space and time at the frequency  $\nu$ ,  $L_\nu$ , in the medium is the RTE (Modest, 2013; Howell *et al.*, 2011) :

$$\left( \frac{n}{c_0} \frac{\partial}{\partial t} + \mathbf{s} \cdot \nabla \right) L_\nu(\mathbf{r}, \mathbf{s}, t) = -(\kappa(\mathbf{r}) + \sigma_s(\mathbf{r})) L_\nu(\mathbf{r}, \mathbf{s}, t) + \sigma_s(\mathbf{r}) \int_{\mathcal{S}^2} L_\nu(\mathbf{r}, \mathbf{s}', t) \Psi(\mathbf{s}, \mathbf{s}') d\mathbf{s}' + S_\nu(\mathbf{r}, \mathbf{s}, t) \quad \forall \mathbf{r} \in \Omega \quad (4.1)$$

where  $L_\nu : (\Omega \times \mathcal{S}^2 \times \mathbb{R}) \mapsto \mathbb{R}$ ,  $t$  is the time variable,  $\mathbf{r} = (r_1, r_2, r_3)$  is the space variable,  $\mathbf{s}$  is the direction,  $n$  is the refractive index (constant),  $c_0$  is the speed of light in vacuum,  $\kappa, \sigma_s : \Omega \mapsto \mathbb{R}^+$  are the absorption and scattering coefficients, respectively,  $\Psi$  is the scattering phase function, which describes the probability that a ray from the direction  $\mathbf{s}'$  will be scattered into the direction  $\mathbf{s}$ , and  $S_\nu$  represents volumetric source terms such as the spontaneous emission.

The RTE is an integro-differential equation that requires heavy computation to get accurate solutions. Thus, an estimation of radiative property maps based on this forward model leads to a very cumbersome inverse problem to solve. The DOT problem consists in using the DA of the RTE as the forward model of the inverse problem (Arridge, 1999; Tarvainen, 2006) :

$$\begin{aligned} \frac{n}{c_0} \frac{\partial \varphi}{\partial t}(\mathbf{r}, t) - \nabla \cdot [D(\mathbf{r}) \nabla \varphi(\mathbf{r}, t)] + \kappa(\mathbf{r}) \varphi(\mathbf{r}, t) &= 0 \quad \forall \mathbf{r} \in \Omega \\ \varphi(\zeta, t) + \frac{A}{2\gamma_{n_\Omega}} D(\zeta) \nabla \varphi(\zeta, t) \cdot \mathbf{n} &= \frac{f(\zeta, t)}{\gamma_{n_\Omega}} \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (4.2)$$

where  $\varphi(\mathbf{r}, t) = \int_{\mathcal{S}^2} L_\nu(\mathbf{r}, \mathbf{s}, t) d\mathbf{s}$  is the photon density,  $\varphi : (\Omega \times \mathbb{R}) \mapsto \mathbb{R}$ ,  $D(\mathbf{r}) = (n_\Omega(\kappa + \sigma))^{-1}$  is the macroscopic scattering coefficient,  $n_\Omega$  is the dimension of  $\Omega$ ,  $\sigma = (1 - g)\sigma_s$  is the reduced scattering coefficient,  $g$  is the asymmetry factor of scattering, which equals the average cosine of the scattering angle,  $\mathbf{n}$  is the unit normal vector to the surface  $\partial\Omega$ ,  $\gamma_{n_\Omega}$  is a fixed parameter dependent on  $n_\Omega$ ,  $A$  is a parameter which characterizes the reflection at the frontier of  $\Omega$  and  $f$  is a diffuse source.

When derivating the DA from the RTE, it is shown that the DA model is a good approximation of the RTE as soon as the medium is highly scattering and satisfies  $0 \ll \kappa \ll \sigma$  (Arridge, 1999; Klose *et al.*, 2002).

Next, using amplitude modulated diffuse source at the frequency  $\omega$ ,  $f_\omega$ , the DA model in the frequency domain writes :

$$\begin{aligned} -\nabla \cdot [D(\mathbf{r})\nabla\varphi(\mathbf{r})] + \left[ \kappa(\mathbf{r}) + \frac{2\pi i \omega n}{c_0} \right] \varphi(\mathbf{r}) &= 0 \quad \forall \mathbf{r} \in \Omega \\ \varphi(\zeta) + \frac{A}{2\gamma_{n_\Omega}} D(\zeta) \nabla\varphi(\zeta) \cdot \mathbf{n} &= \frac{f_\omega(\zeta)}{\gamma_{n_\Omega}} \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (4.3)$$

where  $\varphi : \Omega \mapsto \mathbb{C}$ .

Equation (4.3) leads to compute the state variable  $\varphi$  that depends on space-dependent radiative properties  $\kappa$ ,  $\sigma$ , and which values at the nodes of the sensors will be called predictions. Such modeling is the forward model used in the remainder of this paper which is known to be well-posed mathematically, for example using the complex version of the Lax-Milgram theorem (Brattka et Yoshikawa, 2006).

### 4.3 Cost function and its gradients

Let  $\partial\Omega_d = \partial\Omega_d^1 \cup \dots \cup \partial\Omega_d^L$  be a set of  $L$  surfacic sensors on  $\partial\Omega$  and let us consider  $K$  diffuse sources at different locations along  $\partial\Omega$ . For each source, which constitutes a test, let  $\check{\varphi}_i : \Omega \mapsto \mathbb{C}$  be equal to measurements of the amplitude and phase shift of the transmitted photon density. Predictions  $\varphi_i$  from the forward model (4.3) and measurements  $\check{\varphi}_i$  are integrated to the cost function to be minimized :

$$j(\kappa, \sigma) = \frac{1}{2} \sum_{i=1}^K \int_{\partial\Omega_d} \left| \frac{\varphi_i - \check{\varphi}_i}{\check{\varphi}_i} \right|^2 d\zeta \quad (4.4)$$

The inverse problem can then be formulated as a constrained optimization problem of the form :

$$\text{Find } (\bar{\kappa}, \bar{\sigma}) = \arg \min_{\kappa, \sigma} j(\kappa, \sigma) \text{ subject to } \varphi_i \text{ solution of (4.3), } \forall i \quad (4.5)$$

As a gradient-based method will be used to solve the optimization problem (4.5), gradients of the cost function relative to  $\kappa$  and  $\sigma$  have to be computed. For such a large scale optimization problem, the gradient has to be computed, for CPU-time considerations, by the very fast adjoint-state method. It can be shown that (Arridge, 1999) :

$$\begin{aligned} \nabla_{\kappa} j &= \sum_{i=1}^K \text{Re} (\varphi_i \bar{\varphi}_i^* - n_{\Omega} D^2 \nabla \varphi_i \cdot \nabla \bar{\varphi}_i^*) \\ \nabla_{\sigma} j &= \sum_{i=1}^K \text{Re} (-n_{\Omega} D^2 \nabla \varphi_i \cdot \nabla \bar{\varphi}_i^*) \end{aligned} \quad (4.6)$$

where  $\text{Re}(u)$  and  $\bar{u}$  stand for the real part and complex conjugate of the expression  $u$ , respectively, and  $\varphi_i^*$  is the adjoint variable of the state variable  $\varphi_i$ ,  $\varphi_i^*$  being solution of the adjoint model :

$$\begin{aligned} -\nabla \cdot (D(\mathbf{r}) \nabla \varphi_i^*(\mathbf{r})) + (\kappa(\mathbf{r}) - \frac{2\pi i \omega n}{c_0}) \varphi_i^*(\mathbf{r}) &= 0 \quad \forall \mathbf{r} \in \Omega \\ \frac{2\gamma_{n\Omega}}{A} \varphi_i^*(\zeta) + D(\zeta) \nabla \varphi_i^*(\zeta) \cdot \mathbf{n} &= -\frac{1}{|\check{\varphi}_i(\zeta)|^2} (\varphi_i(\zeta) - \check{\varphi}_i(\zeta)) \mathbb{1}_{[\zeta \in \partial\Omega_d]} \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (4.7)$$

where  $\mathbb{1}_{[\cdot]}$  is the indicator function. In this paper, the forward and adjoint models (4.3)-(4.7) will be solved by the finite element method using the  $\mathcal{P}_1$ -finite element discretization. The reader is invited to consult (Arridge et Schweiger, 1995; Arridge, 1999) for details concerning variational formulations and associated matrix systems to solve (4.3)-(4.7) with the finite element method.

The two radiative properties to be retrieved from data are different in nature, and their order of magnitude also differs. As a consequence, the cost function gradient parts associated with both radiative properties also differ by roughly the same order of magnitude, which is very bad for the convergence. In order to speed-up the iterative convergence to the local minimum,

the proposed strategy consists in performing a scaling on the parameters at the beginning of the optimization problem. Choosing an a priori function for both radiative properties, say  $\kappa_{ap}$  and  $\sigma_{ap}$ , one searches parameters that fluctuate about unity. This adimensionalization leads to recover both  $\varkappa(\mathbf{r}) = \kappa(\mathbf{r})/\kappa_{ap}(\mathbf{r})$  and  $\varsigma(\mathbf{r}) = \sigma(\mathbf{r})/\sigma_{ap}(\mathbf{r})$  for which magnitude is of order one approximately for both coefficients. It results that the considered cost function becomes  $j(\varkappa, \varsigma)$  and the gradients become :

$$\nabla_{\varkappa} j = \kappa_{ap} \nabla_{\kappa} j \quad ; \quad \nabla_{\varsigma} j = \sigma_{ap} \nabla_{\sigma} j \quad (4.8)$$

#### 4.4 Optimization

Although lots of algorithms solving DOT problems in the literature are based on the GN method (Schweiger *et al.*, 2005; Niu *et al.*, 2008; Dehghani *et al.*, 2009b; Tarvainen *et al.*, 2011), the L-BFGS algorithm is used in this paper. The rationale behind this choice was shown in a previous work by the authors (Favennec *et al.*, 2013) : compared to the GN algorithm, the L-BFGS yields much better reconstructions at lower computational price and is much less sensitive to the ill-posed character of the DOT problem. Indeed, in our applications, regularization may be not compulsory when using the L-BFGS algorithm, although its use may however enhance regularity of the solutions. However, penalization type regularization is absolutely compulsory when considering optimization algorithms that rely on matrix inversion, such as with the GN method.

Generally speaking, two ingredients are important in gradient methods : the descent direction and the step-size. At the iteration  $m$ , the descent direction  $d_m$  of the L-BFGS method is given by :

$$d_m = -H_m^{-1} g_m \quad (4.9)$$

where  $g_m$  is the gradient of the cost function and  $H_m^{-1}$  is computed by the update formula :

$$H_{m+1}^{-1} = H_m^{-1} + \frac{s_m s_m^\top}{y_m^\top s_m} \left[ \frac{y_m^\top H_m^{-1} y_m}{y_m^\top s_m} + I_d \right] - \frac{1}{y_m^\top s_m} \left[ s_m y_m^\top H_m^{-1} + H_m^{-1} y_m s_m^\top \right] \quad (4.10)$$

where  $s_m = (\kappa_{m+1} - \kappa_m, \sigma_{m+1} - \sigma_m)$ ,  $y_m = g_{m+1} - g_m$  and  $H_0$  is a positive-definite matrix (identity matrix in this study). According to (Nocedal, 1980), the matrix  $H_m^{-1}$  shall not explicitly be computed in the algorithm for the sake of efficiency. Vectors  $s_m$  and  $y_m$  are stored during the iterations and allow the computation of the direction update (4.9).

Being given a descent direction  $d$ , the step-size is computed with the help of a line-search method which consists in solving the one-dimensional optimization problem :

$$\alpha^* \approx \arg \min_{\alpha > 0} j((\kappa, \sigma) + \alpha d) \quad (4.11)$$

A comparative study for the choice of the line-search with the L-BFGS method has been realized and presented in (Dubot *et al.*, 2014). The golden-section search (algorithm 4.2 in (Antoniou et Lu, 2007)), the quadratic interpolation search (algorithm 4.3 in (Antoniou et Lu, 2007)) and an inexact line-search due to Fletcher (Fletcher, 1980, p. 126) have been considered. The conclusion is that the quadratic interpolation method is the most efficient, in the proposed DOT applications, among the three tested methods. The implementation of the L-BFGS method used in this paper is presented in Algorithm 4.1. A maximum iteration number  $N_M$  is given in order to ensure that the algorithm stops after a finite number of iterations. It has been observed that a maximum iteration number of a few hundred is a good choice when the algorithm does not end by other stopping criteria ( $N_M = 300$  in this paper). In particular, if an higher maximum iteration number is given, the cost function could continue to decrease but the quality of reconstructions will deteriorate. With gradient-based methods such as the L-BFGS, the iteration number plays the role of regularization in the context of Alifanov's iterative regularization (Alifanov, 1974). Concerning initial guesses of the radiative properties  $\gamma^0 = (\varkappa^0, \varsigma^0)$ , they are generally determined by a fitting algorithm applied to the data averaged over all sources making the assumption of a homogeneous medium (McBride *et al.*, 2001a). In this paper, several choices of guessed radiative properties will be systematically considered in the numerical results section 4.6.

**Algorithme 4.1 : The L-BFGS algorithm.**

```

Input :  $k \leftarrow 1, \gamma^0 = (\mathcal{X}^0, \zeta^0), N_M, \varepsilon_1, \varepsilon_2;$ 

Compute  $j(\gamma^0), \nabla j(\gamma^0);$ 
 $d^0 \leftarrow -\nabla j(\gamma^0);$ 
 $\alpha \leftarrow \text{Line-Search}(\gamma^0, d^0, j(\gamma^0));$ 
 $\gamma^1 \leftarrow \gamma^0 + \alpha d^0;$ 
Compute  $j(\gamma^1), \nabla j(\gamma^1);$ 
 $s(0) \leftarrow \gamma^1 - \gamma^0; y(0) \leftarrow \nabla j(\gamma^1) - \nabla j(\gamma^0);$ 
while  $k \leq N_M$  and  $\frac{\|\nabla j(\gamma^k)\|}{\|\nabla j(\gamma^0)\|} > \varepsilon_1$  and  $\frac{|j(\gamma^k) - j(\gamma^{k-1})|}{j(\gamma^k)} > \varepsilon_2$  do
     $d^k \leftarrow \text{Direction-Update}(s, y, \nabla j(\gamma^k), k)$  following (Nocedal, 1980);
     $\alpha \leftarrow \text{Line-Search}(\gamma^k, d^k, j(\gamma^k));$ 
     $\gamma^{k+1} \leftarrow \gamma^k + \alpha d^k;$ 
    Compute  $j(\gamma^{k+1}), \nabla j(\gamma^{k+1});$ 
     $s(k) \leftarrow \gamma^{k+1} - \gamma^k; y(k) \leftarrow \nabla j(\gamma^{k+1}) - \nabla j(\gamma^k);$ 
     $k \leftarrow k + 1;$ 
end
return  $\gamma^k;$ 

```

## 4.5 Wavelet multi-scale method

In the first part of this section, an introduction to the wavelet theory is provided from (Mallat, 1999). Then, in the second part, the strategy used for the development of a wavelet multi-scale method in the context of the proposed inverse problem is presented.

### 4.5.1 Orthogonal wavelet bases

**Definition 1.** *The sequence of closed subspaces of  $L^2(\mathbb{R})$ ,  $\{V_j\}_{j \in \mathbb{Z}}$ , is a multiresolution approximation if the following 6 properties are satisfied (Mallat, 1999) :*

- $\forall (j, k) \in \mathbb{Z}^2, f(r_1) \in V_j \iff f(r_1 - 2^j k) \in V_j$
- $\forall j \in \mathbb{Z}, V_{j+1} \subset V_j$
- $\forall j \in \mathbb{Z}, f(r_1) \in V_j \iff f(r_1/2) \in V_{j+1}$
- $\lim_{j \rightarrow +\infty} V_j = \cap_{j=-\infty}^{+\infty} V_j = \{0\}$

$$e. \quad \lim_{j \rightarrow -\infty} V_j = \overline{\cup_{j=-\infty}^{+\infty} V_j} = L^2(\mathbb{R})$$

*f. There exists  $\theta$  such that  $\{\theta(r_1 - n)\}_{n \in \mathbb{Z}}$  is a Riesz basis of  $V_0$ .*

The approximation of  $f$  at the scale  $2^j$  is defined as the orthogonal projection  $P_{V_j}f$  on  $V_j$ . To compute this projection, one has to find an orthonormal basis of  $V_j$  for all  $j$ . The following theorem constructs this basis from the Riesz basis  $\{\theta(r_1 - n)\}_{n \in \mathbb{Z}}$  (Mallat, 1999) :

**Theorem 1.** *Let  $\{V_j\}_{j \in \mathbb{Z}}$  be a multiresolution approximation and  $\Phi$  be the scaling function whose Fourier transform is*

$$\hat{\Phi}(w) = \frac{\hat{\theta}(w)}{\left(\sum_{k=-\infty}^{+\infty} |\hat{\theta}(w + 2k\pi)|^2\right)^{1/2}} \quad (4.12)$$

*Then, the family*

$$\left\{ \Phi_{j,n}(r_1) = 2^{-j/2} \Phi(2^{-j}(r_1 - n)) \right\}_{n \in \mathbb{Z}} \quad (4.13)$$

*is an orthonormal basis of  $V_j$  for all  $j \in \mathbb{Z}$ .*

The conjugate mirror filter  $h$ , which plays an important role in the fast wavelet transform algorithm, is now defined, with the inner product on  $L^2(\mathbb{R})$  noted  $\langle \cdot, \cdot \rangle$  :

**Definition 2.** *Let  $\Phi \in L^2(\mathbb{R})$  be an integrable scaling function. The conjugate mirror filter  $h$  is defined as*

$$h[n] = \langle 2^{-1/2} \Phi(r_1/2), \Phi(r_1 - n) \rangle \quad (4.14)$$

*The sequence  $h[n]$  may be interpreted as a lowpass discrete filter which smooths the data.*

The approximations of  $f$  at scales  $2^j$  and  $2^{j-1}$  are equal to their orthogonal projections on  $V_j$  and  $V_{j-1}$ , respectively. Let us note  $P_{V_j}f$  and  $P_{V_{j-1}}f$  as these projections. As  $V_j \subset V_{j-1}$ , the

orthogonal complement  $W_j$  of  $V_j$  in  $V_{j-1}$  can be considered. Then we have :

$$V_{j-1} = V_j \oplus W_j \quad (4.15)$$

$$\Rightarrow \forall f \in L^2(\mathbb{R}), \quad P_{V_{j-1}}f = P_{V_j}f + P_{W_j}f \quad (4.16)$$

$$\sum_{n=-\infty}^{+\infty} \langle f, \Phi_{j-1,n} \rangle \Phi_{j-1,n} = \sum_{n=-\infty}^{+\infty} \langle f, \Phi_{j,n} \rangle \Phi_{j,n} + \sum_{n=-\infty}^{+\infty} \langle f, \Psi_{j,n} \rangle \Psi_{j,n} \quad (4.17)$$

$$\sum_{n=-\infty}^{+\infty} a_{j-1}[n] \Phi_{j-1,n} = \sum_{n=-\infty}^{+\infty} a_j[n] \Phi_{j,n} + \sum_{n=-\infty}^{+\infty} d_j[n] \Psi_{j,n} \quad (4.18)$$

where  $\{\Psi_{j,n}(r_1)\}_{n \in \mathbb{Z}}$  has to be an orthonormal basis of  $W_j$  for all  $j \in \mathbb{Z}$ . As written in (Mallat, 1999) : “the complement  $P_{W_j}f$  provides the “details” of  $f$  that appear at the scale  $2^{j-1}$  but which disappear at the coarser scale  $2^j$ ”. The following theorem due to Mallat and Meyer proves that one can construct an orthonormal basis of  $W_j$  by scaling and translating a wavelet  $\Psi$  (Mallat, 1999).

**Theorem 2.** *Let  $\Phi$  be a scaling function and  $h$  the corresponding conjugate mirror filter. Let  $\Psi$  be the function whose Fourier transform is*

$$\hat{\Psi}(w) = 2^{-1/2} \hat{g}(w/2) \hat{\Phi}(w/2) \quad (4.19)$$

with

$$\hat{g}(w) = e^{-iw} \bar{\hat{h}}(w + \pi) \quad (4.20)$$

Let us denote

$$\Psi_{j,n}(r_1) = 2^{-j/2} \Psi(2^{-j}(r_1 - 2^j n)) \quad (4.21)$$

For any scale  $2^j$ ,  $\{\Psi_{j,n}\}_{n \in \mathbb{Z}}$  is an orthonormal basis of  $W_j$ . For all scales,  $\{\Psi_{j,n}\}_{(j,n) \in \mathbb{Z}^2}$  is an orthonormal basis of  $L^2(\mathbb{R})$ .

It can be shown that the inverse Fourier transform of  $\hat{g}$  (see (4.20)) is equal to :

$$g[n] = \langle 2^{-1/2} \Psi(t/2), \Phi(r_1 - n) \rangle \quad (4.22)$$

$$= (-1)^{1-n} h[1-n] \quad (4.23)$$



The sequence  $g$  is the conjugate mirror filter associated to the wavelet  $\Psi$  and may physically be interpreted as a highpass discrete filter which picks out the details of the signal.

Finally, the Mallat's fast orthogonal wavelet transform is given, which allows the fast computation of scaling and wavelet coefficient vectors  $a_j$  and  $d_j$  (see (4.18)), respectively, from  $a_{j-1}$  (Mallat, 1999).

**Theorem 3.** *At the decomposition, coefficients are computed such that :*

$$a_{j+1}[p] = \sum_{n \in \mathbb{Z}} h[n-2p]a_j[n] \quad (4.24)$$

$$d_{j+1}[p] = \sum_{n \in \mathbb{Z}} g[n-2p]a_j[n] \quad (4.25)$$

*At the reconstruction, coefficients are computed such that :*

$$a_j[p] = \sum_{n \in \mathbb{Z}} h[p-2n]a_{j+1}[n] + \sum_{n \in \mathbb{Z}} g[p-2n]d_{j+1}[n] \quad (4.26)$$

Although the fast bi-dimensional orthogonal wavelet transform will be used in the multi-scale reconstruction method, the reader is invited to consult (Mallat, 1999) to obtain explanations about wavelets in higher dimensions. Generally speaking, multiresolutions in higher dimensions are direct extensions of the one-dimensional case using tensor product spaces.

#### 4.5.2 Multi-scale reconstruction method

For the sake of simplicity, the wavelet multi-scale method is presented : (i) in the bidimensional case, (ii) with the domain  $\Omega$  discretized on a grid of  $2^{-N} \times 2^{-N}$  points,  $N \leq 0$ , (iii) with each source and sensor constituted of  $2^{-M}$  pointwise measurements on the grid points,  $0 \geq M \geq N$ , and (iv) with a  $\mathcal{P}_1$ -finite element discretization for all functions involved in the inverse method in order to have an obvious correspondence between function values on grid points and their corresponding scale coefficients. Indeed, because a subsampling of a factor 2 is required when applying equations (4.24) and (4.25), Mallat's algorithm, Theorem 3, is designed for signal whose length is an integer power of two. When it is not the case, algorithms exist to apply the

fast wavelet transform, e.g. by adding zeros to the signal so that its length is increased to the next integer power of two (Addison, 2010).

At the beginning of the algorithm, the sources and experimental measurements (or synthetic data) are decomposed on the smallest scale  $2^M$  (i.e. belong to the  $V_M$  space). This can be achieved by making a correspondence between localizations of grid points in the mesh and scale coefficients in scale coefficient vectors, and setting the scale coefficients to the values of the functions multiplied by  $2^{M/2}$  (Mallat, 1999). Similarly, the initial radiative properties are decomposed on the smallest scale  $2^N$  by setting the scale coefficients to the values of the functions multiplied by  $2^N$  (due to the dimension two). Inverse multiplications, respectively by  $2^{-M/2}$  and  $2^{-N}$ , are performed to pass from the scale coefficients to the function values at the grid points. The same procedure is considered when the wavelet transform is applied to the forward and adjoint fields.

The wavelet multi-scale method consists in applying the L-BFGS method until convergence, from the largest scale ( $2^{N-M}$ ) to the smallest one ( $2^N$ ), with, at a given scale, initializing the radiative properties to the solution of the optimizer at the upper scale. As the two main ingredients of the L-BFGS algorithm are the computation of the cost function and its gradient, the way these functions are computed for a given scale is detailed hereafter in Algorithms 4.2 and 4.3.

Three wavelets are considered in this paper. The first is the Haar wavelet ( $h = [h_0, h_1] = [2^{-1/2}, 2^{-1/2}]$ ), the second is the Daubechies D4 wavelet ( $h = [h_i]_{i=0}^3$ ) and the third is the Daubechies D6 wavelet ( $h = [h_i]_{i=0}^5$ ). The coefficients of the conjugate mirror filters  $h$  for the D4 and D6 wavelets are not given here but can be found in (Mallat, 1999) for instance. It should be noted that when the Haar wavelet is used, no particular attention has to be given when applying the fast wavelet transform. However, an edge problem appears when considering equations (4.24), (4.25) and (4.26) with a finite length sequence  $a_j$  and a wavelet for which the conjugate mirror filter length is greater than two. Among the known and widely used approaches to deal with this problem, let us mention the extension by zeros, the periodic extension, and the sym-

metric extension (Mallat, 1999). Let us note that in all of these methods can occur undesirable edge effects in the transformed signal. These three approaches have been implemented and tested applying the wavelet multi-scale method with the Daubechies D4 and D6 wavelets. It has been observed that the extension by zeros and the periodic extension provide very poor results when compared to the symmetric extension. The latter approach was therefore chosen.

Algorithme 4.2 : Computation of the cost function at scale  $2^{N-l}$ ,  $M \leq l < 0$ .

```

Input : Experimental measurements at scale  $2^{M-l}$ ;

for  $i \leftarrow 1$  to  $K$  do
    Compute the state variable  $\varphi_i$  at the smallest scale  $2^N$ ;
    Extract the complex values of the photon density  $\varphi_i$  at the nodes of the sensors, for
    each sensor  $\partial\Omega_d^1, \dots, \partial\Omega_d^L$ , to obtain  $L$  vectors of length  $2^{-M}$ ;
    Perform the fast wavelet transform following equation (4.24) from scale  $2^M$  to scale
     $2^{M-l}$ , for each one of the  $L$  vectors previously built;
    Add the contribution of the test  $i$  to the cost function at scale  $2^{N-l}$ , computing
    equation (4.4) with experimental measurements and values of the state variable at
    the nodes of the sensors at the scale  $2^{M-l}$ ;
end
return The value of the cost function at scale  $2^{N-l}$ ;

```

## 4.6 Numerical results

### 4.6.1 Tests presentation

The proposed method has been implemented in the freefem++ environment (Hecht, 2012) and some results are presented in this section. The wavelet multi-scale method is applied and compared with the ordinary L-BFGS method to the reconstruction of two bidimensional radiative property maps,  $\kappa(\mathbf{r})$  and  $\sigma(\mathbf{r})$ , where the domain considered,  $\Omega$ , is a square of 4 cm length

Algorithme 4.3 : Computation of the gradient of the cost function at scale  $2^{N-l}$ ,  $M \leq l < 0$ .

**Input** : Experimental measurements at scale  $2^{M-l}$ ;

**for**  $i \leftarrow 1$  **to**  $K$  **do**

    Compute the state variable  $\varphi_i$  at the smallest scale  $2^N$ ;

    Compute the source term of the adjoint model (right term of equation (4.7)) at scale  $2^{M-l}$ , following an analog procedure of steps 2 and 3 of the cost function computation algorithm;

    Compute the adjoint state variable  $\varphi_i^*$  at the smallest scale  $2^N$ ;

    Perform the fast wavelet transform of  $\varphi_i$ ,  $\varphi_i^*$ ,  $\partial_{r_1}(\varphi_i)$ ,  $\partial_{r_2}(\varphi_i)$ ,  $\partial_{r_1}(\varphi_i^*)$  and  $\partial_{r_2}(\varphi_i^*)$  from scale  $2^N$  to scale  $2^{N-l}$ ;

    Add the contribution of the test  $i$  to the gradient of the cost function at scale  $2^{N-l}$ , computing equation (4.6) with the help of  $\varphi_i$ ,  $\varphi_i^*$ ,  $\partial_{r_1}(\varphi_i)$ ,  $\partial_{r_2}(\varphi_i)$ ,  $\partial_{r_1}(\varphi_i^*)$  and  $\partial_{r_2}(\varphi_i^*)$  previously computed at the scale  $2^{N-l}$ ;

**end**

**return** *The gradient of the cost function at scale  $2^{N-l}$ ;*

( $\Omega = [-2, 2]^2$ ). The target properties to be reconstructed,  $\kappa_t$  and  $\sigma_t$ , are defined by

$$\begin{aligned} \kappa_t(r) &= 0.1 \text{ cm}^{-1}, & \sigma_t(r) &= 30 \text{ cm}^{-1}, & r &\in \Omega_1, & \Omega_1 &= [-1.5, -0.5]^2 \\ \kappa_t(r) &= 0.06 \text{ cm}^{-1}, & \sigma_t(r) &= 10 \text{ cm}^{-1}, & r &\in \Omega_2, & \Omega_2 &= [0.5, 1.5]^2 \\ \kappa_t(r) &= 0.08 \text{ cm}^{-1}, & \sigma_t(r) &= 20 \text{ cm}^{-1}, & r &\in \Omega \setminus (\Omega_1 \cup \Omega_2) \end{aligned} \quad (4.27)$$

for the first test, and by

$$\begin{aligned} \kappa_t(r) &= 0.08 + 0.06 \cos(\sqrt{\frac{r_1^2 + r_2^2}{2}}) \text{ cm}^{-1}, & \sigma_t(r) &= 20 + 15 \cos(\sqrt{\frac{r_1^2 + r_2^2}{2}}) \text{ cm}^{-1}, & r &\in \Omega_3 \\ \kappa_t(r) &= 0.08 \text{ cm}^{-1}, & \sigma_t(r) &= 20 \text{ cm}^{-1}, & r &\in \Omega \setminus \Omega_3 \end{aligned} \quad (4.28)$$

where  $\Omega_3 = \left\{ (r_1, r_2) \in \mathbb{R}^2 \text{ such that } r_1^2 + r_2^2 \leq \left(\frac{\pi}{2}\right)^2 \right\}$ , for the second test. The regular spatial discretization associated to  $\Omega$  is composed of  $128 \times 128$  points ( $N = -7$ ). On each side of the square, one source and two sensors are located. Let us note  $\varepsilon = 1/127$ . Then, with respect to  $\varepsilon$ , the source is located on the interval  $[-2 + 56 \times \varepsilon, -2 + 71 \times \varepsilon]$  while the two sensors are respectively located on intervals  $[-2 + 21 \times \varepsilon, -2 + 36 \times \varepsilon]$  and  $[-2 + 91 \times \varepsilon, -2 + 106 \times \varepsilon]$  ( $M = -4$ ).

The physical parameters involved in (4.3) are fixed to :  $n = 1.4$ ,  $\omega = 100 \times 10^6 \text{ s}^{-1}$ ,  $c_0 = 3 \times 10^8 \text{ m.s}^{-1}$ ,  $f_\omega = 1 \text{ W.m}^{-2}$ ,  $\gamma = \pi^{-1}$  and  $A$  is derived from Fresnel's law (see (Dehghani *et al.*, 2009b)). Synthetic data are considered for these numerical tests. These data,  $\check{\phi}$ , representing the pseudo-experimental measurements, are built on a finer mesh ( $255 \times 255$  points) than that of generating the predictions  $\phi$  in order to avoid the inverse crime (Colton et Kress, 1992; Kaipio et Somersalo, 2005). Then, a multiplicative white gaussian noise is applied to  $\check{\phi}$  at the nodes of the sensors to simulate the noise inherent to experimental devices, i.e  $\check{\phi}_{\text{noisy}} = \check{\phi}(1 + 0.01 \times m)$  where  $m \sim \mathcal{N}(0, 1)$ .

The fast wavelet transform is applied with the Haar, D4 or D6 wavelets to decompose the sources and synthetic data at scales  $-3, \dots, 0$  and thus, to obtain a series of inverse problems for scales  $-7, \dots, -3$  following Algorithms 4.2 and 4.3. Concerning the comparison of the wavelet multi-scale method and the ordinary L-BFGS method, the choice was made to allow the same total number of iterations for the two methods. Moreover, the larger the scale, the faster the convergence of the solution toward the local minimum, and thus the allowed maximal number of iterations will be reduced. Therefore, the inputs of the ordinary L-BFGS method, Algorithm 4.1, are the following :  $N_M = 300$ ,  $\varepsilon_1 = 10^{-4}$  and  $\varepsilon_2 = 10^{-8}$ . The same inputs are considered for the wavelet multi-scale method, except the maximum number of iterations which is fixed to  $N_M = 20, 40, 60, 80$  and  $100$  for scales  $-3, -4, -5, -6$  and  $7$ , respectively.

#### 4.6.2 Definition of reconstruction errors

In order to gauge accuracy of the reconstructions, two errors are introduced. Let  $\lambda^j$  and  $\lambda_t^j$  be either the reconstructed and target absorption or reduced scattering coefficients at the node  $j$ , respectively, and  $N_s = 128 \times 128$  be the dimension of the state and control spaces. The first error is the deviation factor,  $e_1^\lambda$ , which measures the deviation of the reconstructed image from

the target image and is defined as :

$$e_1^\lambda = 100 \times \frac{\sqrt{\sum_{j=1}^{N_s} (\lambda^j - \lambda_t^j)^2}}{\sqrt{\sum_{j=1}^{N_s} (\lambda_t^j)^2}} \quad (4.29)$$

A small value of  $e_1^\lambda$  indicates a reconstructed image with high accuracy. The second error is the correlation coefficient (Flannery *et al.*, 1992; Klose et Hielscher, 2003),  $e_2^\lambda$ , which measures the linear correlation between the target and the reconstructed image and is defined as :

$$e_2^\lambda = \frac{\sum_{j=1}^{N_s} (\lambda^j - \bar{\lambda}) (\lambda_t^j - \bar{\lambda}_t)}{(N_s - 1) \sigma_\lambda \sigma_{\lambda_t}} \quad (4.30)$$

where  $\bar{\lambda}$  and  $\bar{\lambda}_t$  are the mean values of vectors  $\lambda$  and  $\lambda_t$  and  $\sigma_\lambda$ ,  $\sigma_{\lambda_t}$  are standard deviations of the target and reconstructed images, respectively, given by

$$\sigma_\lambda = \sqrt{\frac{\sum_{j=1}^{N_s} (\lambda^j - \bar{\lambda})^2}{N_s - 1}} \quad (4.31)$$

and

$$\sigma_{\lambda_t} = \sqrt{\frac{\sum_{j=1}^{N_s} (\lambda_t^j - \bar{\lambda}_t)^2}{N_s - 1}} \quad (4.32)$$

A large value of  $e_2^\lambda$  shows a high detectability in the reconstructed image and indicates a reconstructed image with high accuracy.

#### 4.6.3 Test 1 : reconstruction of discontinuous target properties

For this first test, several reconstructions of target properties, eq. (4.27), are investigated depending on initial guesses for the radiative properties. The control variables  $\varkappa(\mathbf{r})$  and  $\varsigma(\mathbf{r})$  are initialized to unity while values taken by a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  are fixed to those of the background and in the range of more or less ten percent of those of the background. Tables 4.1, 4.2, 4.3 and 4.4 give errors in matrix form  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$ , for all considered initial guesses, obtained

at the end of the inversion with the ordinary L-BFGS method or the wavelet multi-scale method using the Haar, D4 or D6 wavelets. In each table the values of  $\kappa_{ap}$  are in the first column (0.072-0.088) and those for  $\sigma_{ap}$  are along the first row (18-22). The other cells contain errors. Results show that the wavelet multi-scale method using the Haar wavelet outperforms all of the other methods, errors  $e_1^\kappa$  and  $e_1^\sigma$  being smaller and errors  $e_2^\kappa$  and  $e_2^\sigma$  being larger for all the investigated cases. Overall, it is observed that the wavelet multi-scale method using the D6 wavelet outperforms the ordinary L-BFGS method while the wavelet multi-scale method using the D4 wavelet outperforms most of the time the wavelet multi-scale method using the D6 wavelet. Fig. 4.1 and Fig. 4.2 present the target properties and the obtained reconstructions with the four proposed inverse methods for initial guesses  $\kappa_{ap} = 0.08$  and  $\sigma_{ap} = 20$ . It is seen that the wavelet multi-scale method using the Haar wavelet allows to access better reconstructions compared to the three other methods. Specifically, two major improvements should be noted. Firstly, it is observed that the wavelet multi-scale method using the Haar wavelet reduces significantly the presence of inclusions at the edge of the domain  $\Omega$  that have no reason to exist, unlike what is found for the ordinary L-BFGS method (Fig. 4.1-(d)). Secondly, one notes that the location and shape of inclusions are more accurately determined, unlike images obtained with the ordinary L-BFGS method, which, as in classical approaches, are blurred (Cao *et al.*, 2007; Arridge *et al.*, 2008). These improvements are also observed to a lesser degree from reconstructions based on the wavelet multi-scale method using the D4 and D6 wavelets when compared to those of the ordinary L-BFGS method.

TAB. 4.1 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 1 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the ordinary L-BFGS method.

$\kappa_{ap} \setminus \sigma_{ap}$	18	20	22
<b>0.072</b>	$\begin{pmatrix} 7.37 & 12.92 \\ 0.73 & 0.71 \end{pmatrix}$	$\begin{pmatrix} 7.38 & 11.65 \\ 0.68 & 0.75 \end{pmatrix}$	$\begin{pmatrix} 7.48 & 11.31 \\ 0.63 & 0.77 \end{pmatrix}$
<b>0.08</b>	$\begin{pmatrix} 5.27 & 12.12 \\ 0.80 & 0.74 \end{pmatrix}$	$\begin{pmatrix} 5.23 & 10.80 \\ 0.81 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 5.46 & 12.07 \\ 0.80 & 0.74 \end{pmatrix}$
<b>0.088</b>	$\begin{pmatrix} 7.70 & 12.16 \\ 0.74 & 0.75 \end{pmatrix}$	$\begin{pmatrix} 7.64 & 11.70 \\ 0.74 & 0.75 \end{pmatrix}$	$\begin{pmatrix} 7.26 & 12.77 \\ 0.73 & 0.70 \end{pmatrix}$

TAB. 4.2 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 1 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the Haar wavelet.

$\kappa_{ap} \setminus \sigma_{ap}$	18	20	22
<b>0.072</b>	$\begin{pmatrix} 6.02 & 10.64 \\ 0.78 & 0.80 \end{pmatrix}$	$\begin{pmatrix} 6.63 & 9.77 \\ 0.75 & 0.83 \end{pmatrix}$	$\begin{pmatrix} 7.17 & 9.97 \\ 0.71 & 0.82 \end{pmatrix}$
<b>0.08</b>	$\begin{pmatrix} 4.76 & 10.13 \\ 0.84 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 4.77 & 9.61 \\ 0.84 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 4.89 & 9.76 \\ 0.84 & 0.83 \end{pmatrix}$
<b>0.088</b>	$\begin{pmatrix} 6.73 & 9.91 \\ 0.77 & 0.83 \end{pmatrix}$	$\begin{pmatrix} 6.16 & 9.41 \\ 0.78 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 5.92 & 10.13 \\ 0.78 & 0.81 \end{pmatrix}$

TAB. 4.3 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 1 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D4 wavelet.

$\kappa_{ap} \setminus \sigma_{ap}$	18	20	22
<b>0.072</b>	$\begin{pmatrix} 6.47 & 12.24 \\ 0.75 & 0.73 \end{pmatrix}$	$\begin{pmatrix} 7.20 & 10.70 \\ 0.75 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 7.53 & 10.66 \\ 0.71 & 0.80 \end{pmatrix}$
<b>0.08</b>	$\begin{pmatrix} 5.33 & 11.29 \\ 0.81 & 0.78 \end{pmatrix}$	$\begin{pmatrix} 5.02 & 10.06 \\ 0.82 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 5.20 & 10.88 \\ 0.81 & 0.79 \end{pmatrix}$
<b>0.088</b>	$\begin{pmatrix} 7.19 & 10.56 \\ 0.75 & 0.81 \end{pmatrix}$	$\begin{pmatrix} 6.54 & 10.01 \\ 0.74 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 6.21 & 10.80 \\ 0.76 & 0.79 \end{pmatrix}$

TAB. 4.4 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 1 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D6 wavelet.

$\kappa_{ap} \setminus \sigma_{ap}$	18	20	22
<b>0.072</b>	$\begin{pmatrix} 6.47 & 12.18 \\ 0.73 & 0.73 \end{pmatrix}$	$\begin{pmatrix} 6.67 & 10.74 \\ 0.73 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 7.43 & 12.43 \\ 0.71 & 0.74 \end{pmatrix}$
<b>0.08</b>	$\begin{pmatrix} 5.39 & 10.67 \\ 0.80 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 5.25 & 10.75 \\ 0.81 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 5.40 & 11.54 \\ 0.80 & 0.75 \end{pmatrix}$
<b>0.088</b>	$\begin{pmatrix} 7.88 & 10.82 \\ 0.71 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 7.55 & 11.48 \\ 0.73 & 0.76 \end{pmatrix}$	$\begin{pmatrix} 7.34 & 13.88 \\ 0.66 & 0.62 \end{pmatrix}$



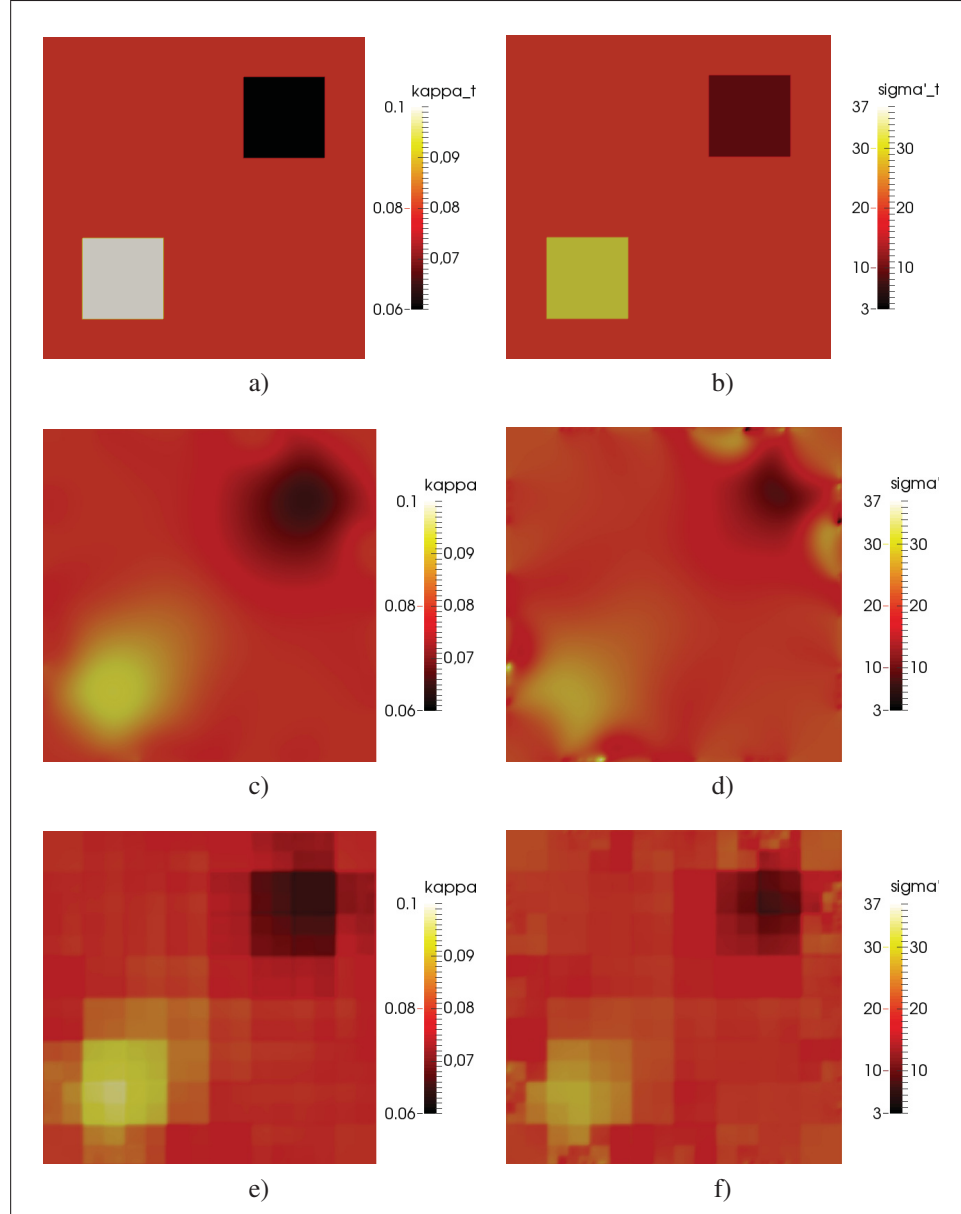


FIG. 4.1 Test 1. Reconstruction of radiative properties with  $\kappa_{ap} = 0.08$  and  $\sigma_{ap} = 20$ . Left-hand side, results for  $\kappa$  ; right-hand side, results for  $\sigma$ . (a-b) target properties ; (c-d) ordinary L-BFGS method ; (e-f) Haar wavelet.

#### 4.6.4 Test 2 : reconstruction of continuous target properties

In this section, in order to more deeply compare the ordinary L-BFGS and wavelet multi-scale methods, reconstructions of continuously varying properties (4.28) which occupy a circle-

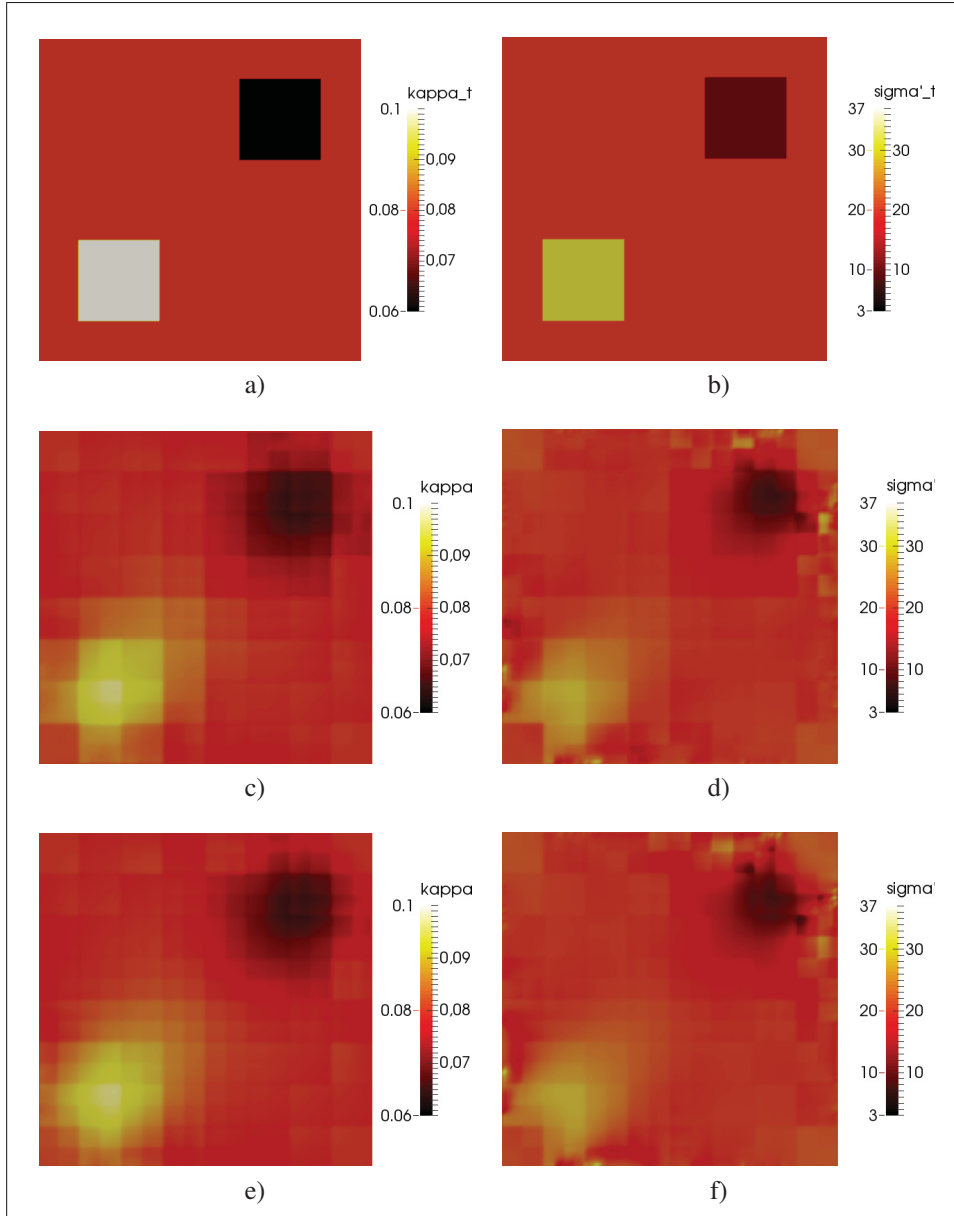


FIG. 4.2 Test 1. Reconstruction of radiative properties with  $\kappa_{ap} = 0.08$  and  $\sigma_{ap} = 20$ . Left-hand side, results for  $\kappa$ ; right-hand side, results for  $\sigma$ . (a-b) target properties; (c-d) D4 wavelet; (e-f) D6 wavelet.

shaped subdomain in  $\Omega$  are investigated. Indeed, wavelet multi-scale methods gave better results for the test 1 partly because the original targets contained rectangular inclusions. But the bidimensional wavelet transform has the property to extract vertical and horizontal edges very well. Especially, because the bidimensional Haar scaling function and the inclusions have

similar shapes. For this second test, a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  are fixed to those of the background ( $\kappa_{ap} = 0.08$ ,  $\sigma_{ap} = 20$ ), to approximately the mean values of target properties ( $\kappa_{ap} = 0.1$ ,  $\sigma_{ap} = 25$ ) or between these two values ( $\kappa_{ap} = 0.09$ ,  $\sigma_{ap} = 22.5$ ). Tables 4.5, 4.6, 4.7 and 4.8 give errors in matrix form, for all considered initial guesses, obtained at the end of the inversion with the ordinary L-BFGS method and the wavelet multi-scale method using the Haar, D4 or D6 wavelets. First, it is observed that all the tests converged for the wavelet multi-scale method using the Haar wavelet while one and five tests diverged for the wavelet multi-scale method using the D4 and D6 wavelets, respectively, and two tests diverged for the ordinary L-BFGS method. Second, it is observed that reconstructions of the reduced scattering coefficient with the wavelet multi-scale method using the Haar wavelet are much more accurate than reconstructions based on the ordinary L-BFGS method. It is also observed that reconstructions of the absorption coefficient with the wavelet multi-scale method using the Haar wavelet are much more accurate than reconstructions based on the ordinary L-BFGS method in terms of the error  $e_1^\kappa$  (except the test  $\kappa_{ap} = 0.09$ ,  $\sigma_{ap} = 22.5$  which gives a close result) but slightly less accurate in terms of the error  $e_2^\kappa$ . It is found that the wavelet multi-scale method using the Haar wavelet clearly outperforms the wavelet multi-scale method using the D4 wavelet and that, overall, the latter gives better results than the wavelet multi-scale method using the D6 wavelet. Fig. 4.3 and Fig. 4.4 present the target properties and the obtained reconstructions with the four proposed inverse methods for initial guesses  $\kappa_{ap} = 0.1$  and  $\sigma_{ap} = 22.5$ . Although this is one of the worst results for reconstructions based on the Haar wavelet, it is observed that the latter gives few false fluctuations at the edge of the domain when compared to the other methods, specially for the reduced scattering coefficient. Fig. 4.5 presents the target properties and reconstructions obtained with the wavelet multi-scale method using the Haar wavelet for the two divergent cases of the ordinary L-BFGS method. It is seen that accurate reconstructions are obtained. Finally, in order to verify that the ordinary L-BFGS method would not give better results if other stopping criteria were chosen, the reconstruction process was stopped just before the divergence of the algorithm for these two cases. For instance, errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix} = \begin{pmatrix} 15.32 & 19.00 \\ 0.92 & 0.53 \end{pmatrix}$  were obtained for the case  $\kappa_{ap} = 0.08$  and  $\sigma_{ap} = 20$ , which in-

icates much less accurate reconstructions than the wavelet multi-scale method using the Haar wavelet.

TAB. 4.5 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 2 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the ordinary L-BFGS method. A cross indicates the divergence of the algorithm.

$\kappa_{ap} \setminus \sigma_{ap}$	20	22.5	25
0.08	×	×	$\begin{pmatrix} 12.58 & 15.10 \\ 0.97 & 0.85 \end{pmatrix}$
0.09	$\begin{pmatrix} 10.59 & 14.13 \\ 0.97 & 0.80 \end{pmatrix}$	$\begin{pmatrix} 7.79 & 12.50 \\ 0.97 & 0.85 \end{pmatrix}$	$\begin{pmatrix} 10.66 & 12.41 \\ 0.97 & 0.86 \end{pmatrix}$
0.1	$\begin{pmatrix} 15.49 & 14.18 \\ 0.97 & 0.85 \end{pmatrix}$	$\begin{pmatrix} 10.03 & 8.80 \\ 0.94 & 0.92 \end{pmatrix}$	$\begin{pmatrix} 11.35 & 10.73 \\ 0.95 & 0.88 \end{pmatrix}$

TAB. 4.6 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 2 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the Haar wavelet.

$\kappa_{ap} \setminus \sigma_{ap}$	20	22.5	25
0.08	$\begin{pmatrix} 6.81 & 7.25 \\ 0.96 & 0.95 \end{pmatrix}$	$\begin{pmatrix} 6.89 & 8.49 \\ 0.96 & 0.94 \end{pmatrix}$	$\begin{pmatrix} 7.14 & 10.98 \\ 0.96 & 0.91 \end{pmatrix}$
0.09	$\begin{pmatrix} 8.17 & 6.84 \\ 0.95 & 0.96 \end{pmatrix}$	$\begin{pmatrix} 7.94 & 7.39 \\ 0.96 & 0.95 \end{pmatrix}$	$\begin{pmatrix} 7.98 & 10.25 \\ 0.96 & 0.91 \end{pmatrix}$
0.1	$\begin{pmatrix} 9.89 & 6.56 \\ 0.92 & 0.96 \end{pmatrix}$	$\begin{pmatrix} 9.88 & 7.60 \\ 0.93 & 0.94 \end{pmatrix}$	$\begin{pmatrix} 9.99 & 9.93 \\ 0.94 & 0.91 \end{pmatrix}$

#### 4.6.5 Discussion

In conclusion, one can state from these two numerical studies that : (i) the wavelet multi-scale method using the Haar wavelet brings more stability with respect to the ordinary L-BFGS method, (ii) the wavelet multi-scale method using the Haar wavelet provides the best reconstructions in most of the tested cases, (iii) the wavelet multi-scale method using the D4 wavelet brings a little more stability with respect to the ordinary L-BFGS method while the wavelet

TAB. 4.7 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 2 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D4 wavelet. A cross indicates the divergence of the algorithm.

$\kappa_{ap} \setminus \sigma_{ap}$	20	22.5	25
<b>0.08</b>	×	$\begin{pmatrix} 8.66 & 14.56 \\ 0.95 & 0.86 \end{pmatrix}$	$\begin{pmatrix} 9.62 & 15.79 \\ 0.95 & 0.86 \end{pmatrix}$
<b>0.09</b>	$\begin{pmatrix} 8.70 & 10.28 \\ 0.95 & 0.90 \end{pmatrix}$	$\begin{pmatrix} 8.52 & 11.35 \\ 0.95 & 0.89 \end{pmatrix}$	$\begin{pmatrix} 8.75 & 12.55 \\ 0.95 & 0.87 \end{pmatrix}$
<b>0.1</b>	$\begin{pmatrix} 11.01 & 9.56 \\ 0.92 & 0.91 \end{pmatrix}$	$\begin{pmatrix} 12.11 & 10.58 \\ 0.94 & 0.89 \end{pmatrix}$	$\begin{pmatrix} 10.94 & 11.38 \\ 0.93 & 0.88 \end{pmatrix}$

TAB. 4.8 Errors  $\begin{pmatrix} e_1^\kappa & e_1^\sigma \\ e_2^\kappa & e_2^\sigma \end{pmatrix}$  for the test 2 and different values of the a priori functions  $\kappa_{ap}$  and  $\sigma_{ap}$  ( $\text{cm}^{-1}$ ) with the wavelet multi-scale method using the D6 wavelet. A cross indicates the divergence of the algorithm.

$\kappa_{ap} \setminus \sigma_{ap}$	20	22.5	25
<b>0.08</b>	×	×	×
<b>0.09</b>	$\begin{pmatrix} 12.55 & 16.03 \\ 0.94 & 0.72 \end{pmatrix}$	$\begin{pmatrix} 11.47 & 14.31 \\ 0.95 & 0.77 \end{pmatrix}$	×
<b>0.1</b>	×	$\begin{pmatrix} 11.51 & 12.36 \\ 0.95 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 12.83 & 13.04 \\ 0.96 & 0.82 \end{pmatrix}$

multi-scale method using the D6 wavelet is the least stable among the four proposed inverse methods, and (iv) the wavelet multi-scale method using the D4 and D6 wavelets does not really enhance the quality of the reconstructions when compared to the ordinary L-BFGS method. Finally, let us note that the reason the wavelet multi-scale method using the D4 and D6 wavelets fails to give better results should come from undesirable edge effects in the transformed sequences and images. Indeed, it is of crucial importance to not exhibit edge effects in the one-dimensional transformed experimental measurements (or synthetic data) and adjoint source terms as well as in the bidimensional transformed forward and adjoint fields, the latter being maximal at the edges of the domain. Thus, it might be interesting to consider a discrete wavelet transform without edge effects using wavelet extrapolation as described in (Williams et Amaratunga, 1997).

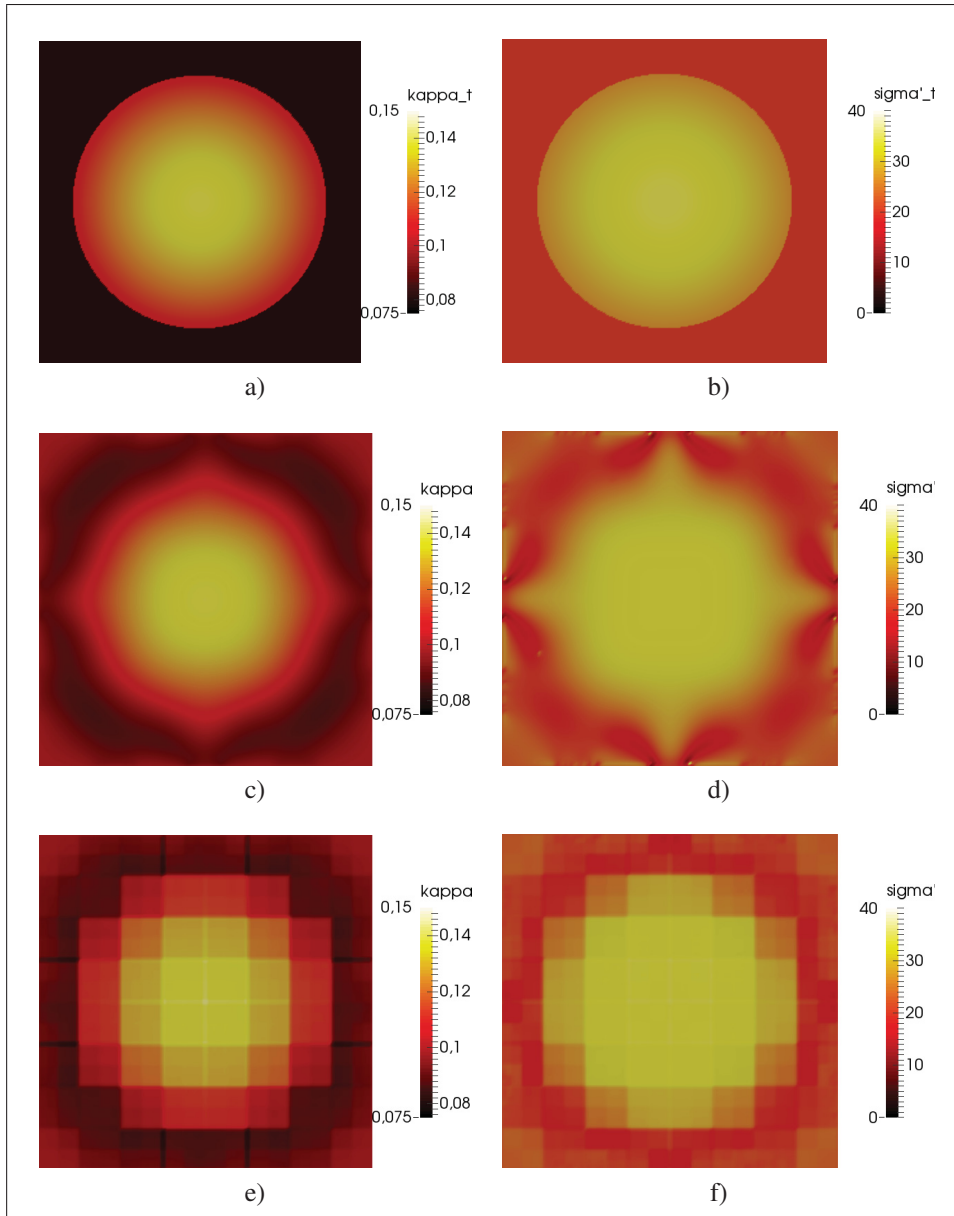


FIG. 4.3 Test 2. Reconstruction of radiative properties with  $\kappa_{ap} = 0.1$  and  $\sigma_{ap} = 22.5$ . Left-hand side, results for  $\kappa$ ; right-hand side, results for  $\sigma$ . (a-b) target properties ; (c-d) ordinary L-BFGS method ; (e-f) Haar wavelet.

## 4.7 Conclusion

A wavelet multi-scale method based on the L-BFGS algorithm using the Haar, Daubechies D4 and D6 wavelets has been designed, implemented and validated for the two-dimensional

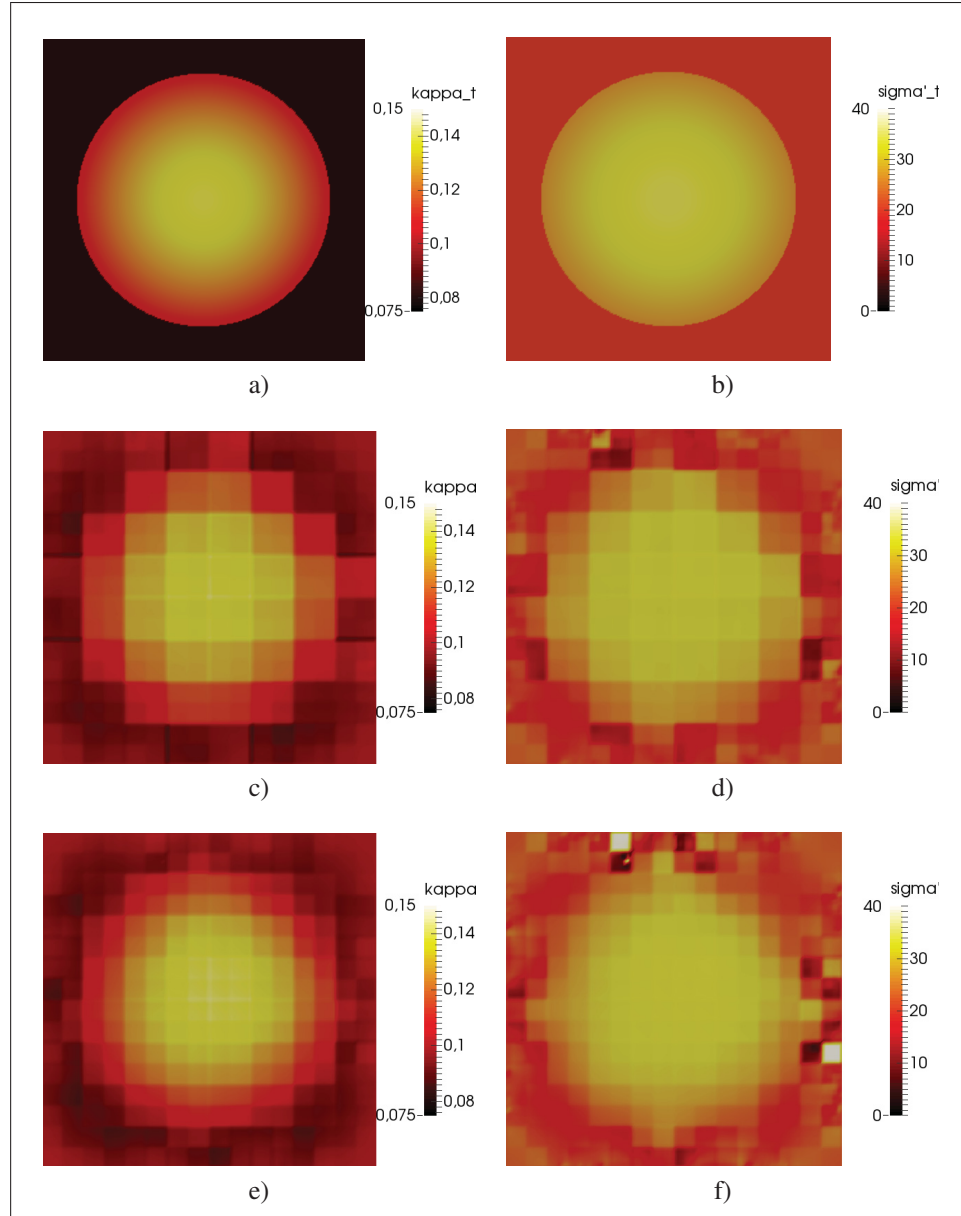


FIG. 4.4 Test 2. Reconstruction of radiative properties with  $\kappa_{ap} = 0.1$  and  $\sigma_{ap} = 22.5$ . Left-hand side, results for  $\kappa$ ; right-hand side, results for  $\sigma$ . (a-b) target properties; (c-d) D4 wavelet; (e-f) D6 wavelet.

inverse problem of diffuse optical tomography in the frequency domain. The numerical results indicate the effectiveness of the method using the Haar wavelet with respect to the ordinary L-BFGS method. In particular, a reduction of the blurring effect has been observed in a numerical test containing rectangular inclusions and a strong stability of the latter method has

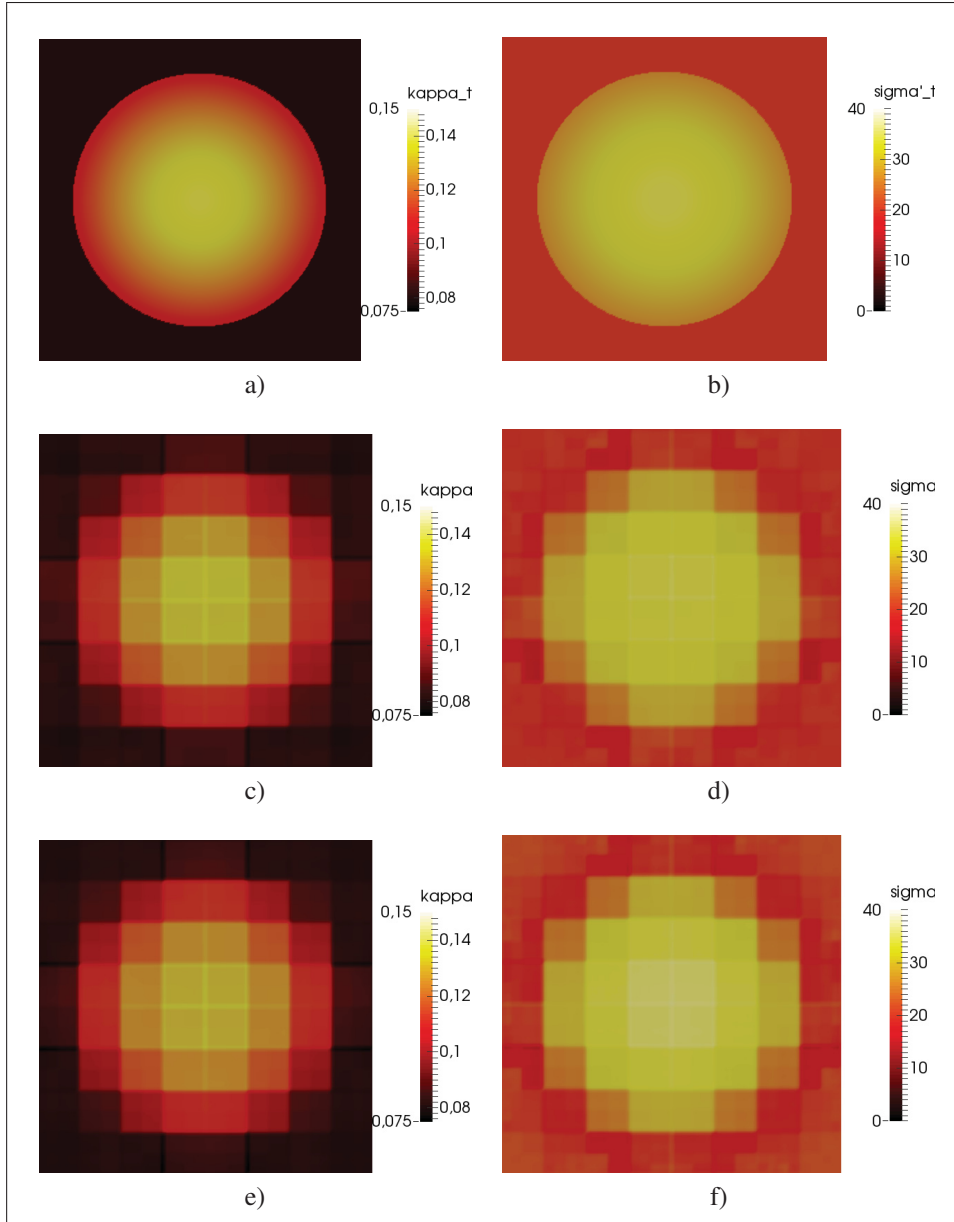


FIG. 4.5 Test 2. Reconstruction of radiative properties with the Haar wavelet. Left-hand side, results for  $\kappa$ ; right-hand side, results for  $\sigma$ . (a-b) target properties; (c-d)  $\kappa_{ap} = 0.08$  and  $\sigma_{ap} = 20$ ; (e-f)  $\kappa_{ap} = 0.08$  and  $\sigma_{ap} = 22.5$ .

been shown on a numerical test containing continuously varying target properties which occupy a circle-shaped subdomain. Upcoming work includes an extension of the method to the three-dimensional case and/or in the time domain, together with the use of a discrete wavelet



transform without edge effects using wavelet extrapolation in order to improve obtained results with the Daubechies D4 and D6 wavelets. The application of the proposed method to the inversion of the full radiative transfer equation is also considered.



## CHAPITRE 5

### ARTICLE 4 : SPACE-DEPENDENT SOBOLEV GRADIENTS AS A REGULARIZATION FOR INVERSE RADIATIVE TRANSFER PROBLEMS

Fabien Dubot<sup>1,2</sup>, Yann Favennec<sup>2</sup>, David Le Hardy<sup>2</sup>, Benoit Rousseau<sup>2</sup>, Daniel R. Rousse<sup>1</sup>

<sup>1</sup> Chaire de recherche industrielle t3e, École de technologie supérieure,

1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

<sup>2</sup> Laboratoire de thermocinétique de Nantes, UMR CNRS 6607,

Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France

Article soumis dans la revue « Journal of Quantitative Spectroscopy & Radiative Transfer »  
en août 2015.

#### Abstract

Diffuse optical tomography problems rely on the solution of an optimization problem for which the dimension of the control space is usually large. Thus, gradient-type optimizers are likely to be used, such as the BFGS algorithm, along with the adjoint-state method to compute the cost function gradient. Usually, the  $L^2$ -inner product is chosen within the extraction procedure while alternative inner products that act as regularization can be used. This paper presents some results based on space-dependent Sobolev inner-products and shows that this method acts as an efficient noise filter. Numerical results indicate that the use of Sobolev gradients can be particularly attractive in the context of inverse problems, particularly because of the simplicity of this regularization.

#### 5.1 Introduction

Diffuse optical tomography (DOT) consists in reconstructing the spatial radiative property distributions of an optically thick participating medium,  $\Omega$ , from optical measurements collected at the frontier of  $\Omega$  (Arridge, 1999; Hielscher *et al.*, 2002; Charette *et al.*, 2008). This inverse problem has been a surge of interest since the middle of the nineties, due to better theoreti-

cal understanding in radiative transfer ; improvements in light detection and control of light sources ; and the increase of the computing resource performances. This optical diagnostic technique is mainly driven by potential applications in biological tissues imaging although it can be used for non-destructive testing in industrial processes, or the characterization of complex semi-transparent media such as cellular foams with spatially varying porosities or fiber reinforced thermoplastic composites with inhomogeneity in the fibers distribution. In this domain, the characterization has often been limited to the reconstruction of constant absorption and scattering coefficients, and anisotropy factor of the phase function, considering the unidimensional radiative transfer equation (RTE) as forward model (Baillis et Sacadura, 2000).

Biomedical applications of the DOT include imaging of finger joints (Hielscher *et al.*, 2004), breast imaging (Pogue *et al.*, 1997; Leff *et al.*, 2008), neuroimaging of humans (Strangman *et al.*, 2002; Boas *et al.*, 2004; Dehghani *et al.*, 2009c) or small-animal imaging (Hielscher, 2005) to name but a few. In this context, near-infrared light is used in order to take advantage of the so-called therapeutic window (600 – 900 nm) in which tissues have low absorption. According to the application, RTE has to be considered as forward model (Klose *et al.*, 2002; Klose et Hielscher, 2003; Ren *et al.*, 2006; Tarvainen *et al.*, 2008a; Balima *et al.*, 2011a, 2013; Asllanaj *et al.*, 2015), diffuse approximation (DA) may be applied (Pogue *et al.*, 1997; Dehghani *et al.*, 2009c; Jiang, 2010; Arridge *et al.*, 2011; Dubot *et al.*, 2015c,b), or hybrid model is employed (Tarvainen *et al.*, 2005b,a, 2008b; Kim et Hielscher, 2010; Tarvainen *et al.*, 2011). This paper deals with two different forward models, namely the bi-dimensional steady-state RTE and the three-dimensional frequency DA.

The evaluation of space-distributed physical properties from the knowledge of incomplete state data is an inverse problem as opposed to forward problems where the state is computed from the knowledge of properties, boundary conditions, etc. Such inverse problems are usually difficult to solve due to their ill-posed nature (Woodbury, 2003).

The solution of the inverse problem is obtained through the minimization of a cost function that measures the misfit between the predictions (solution of the forward problem) and some expe-

rimental measurements. Though zero-order optimization methods may be used for relatively similar problems (Kim *et al.*, 2004; Parwani *et al.*, 2013), gradient-type methods are dealt with in this paper because of the large dimension of the parameter space. Due to the ill-posed nature of the inverse problem, large variations about the solution may be found when no regularization is performed.

In order to limit the amplitude of such oscillations, the ordinary Tikhonov approach consists in assuming that the properties are subject to a minimal norm in a given vectorial space which leads to enhancing the *a priori* regularity of the properties (Tikhonov et Arsenin, 1977; Morozov *et al.*, 1984). This approach relies on the penalization of the cost function adding prior information on the properties to be recovered. However, the weight associated to the penalization function must be chosen carefully. Indeed, a too small value for the penalization function leads to a small cost function but with high property variations, while too high values lead to small property variations about priors but with a high resulting cost function value at the optimum. The compromise minimizing both the cost function and the property variations may be performed according to the construction of the so-called L-curve as suggested by Hansen (Hansen et O’Leary, 1993). Even though, for appropriate penalization, the optimal Tikhonov parameter has been found to be quasi-independent of the parameter space dimension (Dubot *et al.*, 2015a), the construction of such curve, searching its convex corner location, is likely to necessitate a large number of inverse problem solutions for plenty of penalization function weights.

Appropriate parameterization for the space-dependent properties may also affect the solution regularity. As a matter of fact, the amount of measurements data is likely to be insufficient to retrieve the large amount of parameters when considering, at the end, the discrete optimization problem on a fine grid. Thus, regularization by reduction of the dimension of the control space could be employed, which consists in searching the physical properties projected on a coarse finite element grid, as suggested for instance in (Chavent, 2010), and applied for optical tomography applications in (Dubot *et al.*, 2015b).

This paper focuses on a different strategy to improve the reconstruction of space-distributed parameters when the data is corrupted by noise. Rather than modifying the cost function with the addition of a penalization term, the proposed strategy filters the effect of the noise, not on the data as it is done for instance when using the mollification method (Woodbury, 2003; Murio, 1993), but on the cost function gradient itself. This means that the optimization is not modified at all since the original cost function remains the same. It also means that the data is not modified : the whole information contained within the experimental data is integrated within the cost function to be minimized. As a matter of fact, several inner products are to be defined when considering the inverse optimization problem. According to Protas (Protas, 2008), each of them (cost function definition, adjoint identity, and gradient extraction) may somehow yield to a particular regularization opportunity. The noise present in the measurements data propagates through the adjoint problem and then to the cost function gradient from which the directions of descent are computed to supply the optimizer. The propagation of the noise induces high variations of the cost function gradient with respect to what the gradient would be without noise. The idea is thus to somehow apply a low-pass filter to the cost function gradient, exactly where the noise is the highest. This is done choosing the space-dependent weighted Sobolev inner product definition when extracting the cost function gradient rather than the ordinary Hilbert inner product. Note that space-independent Sobolev inner products have been used in other situations, in image treatment (*e.g.* (Renka, 2009)), in Partial Differential Equations (PDEs) energy minimization (*e.g.* (Danaila et Kazemi, 2010)), and for boundary conditions reconstruction with accurate final time (Nassiopoulos et Bourquin, 2010; Bourquin et Nassiopoulos, 2011).

The underlying application of this study is optical tomography, i.e. the reconstruction of radiative properties from light intensity measurements on some parts of the boundary. The paper is therefore organized as follows. Section 5.2 gives the theoretical bases for extracting the cost function gradients when the cost function is expressed in terms of a state which is solution of some PDEs. This section is written in a concise way, following Maillet, Favennec and Le Masson (Maillet *et al.*, 2011), such that the results can be easily used for other practical appli-

cations. Also, only the continuous equations are dealt with, for the states, the adjoints and the gradients. This “differentiate-then-discretize” approach has the advantage of versatility when considering afterwards different discretization choices (one for the states, one for the physical properties, etc.). The cost function gradient is then extracted firstly in a general framework using the ordinary  $L^2$ -inner product then using space-independent and finally space-dependent Sobolev inner products. Section 5.3 presents application of space-dependent Sobolev gradients for the inverse problem of optical tomography in which the forward model is the RTE. The continuous cost function gradient is fully derived and numerical tests are performed for several levels of noise. As the Sobolev filter is chosen to follow an exponential function, only one parameter fully describes the regularization. Next, the frequency DA in the three-dimensional case is considered as forward model of light propagation in Section 5.4. Finally, Section 5.5 naturally summarizes the main conclusions and perspectives of the current work. Note that the proposed method has been implemented within the FreeFem++ environment (Hecht, 2012) while the Paraview software (Ayachit, 2015) was used for 2D and 3D results visualization.

## 5.2 Space-dependent Sobolev cost function gradient

Let us define the domain of interest  $\Omega \subset \mathbb{R}^{n_\Omega}$  and denote the cost function  $\mathcal{J}(u)$  expressed in terms of the state  $u : \Omega \rightarrow \mathbb{K}$ . This cost function may be expressed as the half of the squared norm of a quadratic function  $v$  based on the discrepancy between some state predictions  $u$  and the related measurements  $u_d$ . The norm is expressed with a complex inner product  $(\cdot, \cdot)_{\mathcal{X}}$  defined later on for specific applications. Cost function  $\mathcal{J}(u)$  is actually to be minimized with respect to some space-distributed parameters  $\psi \in \mathcal{P}$ , with  $\mathcal{P} \subseteq L^2(\Omega)$  for example (depending on the forward model considered). The generic cost function is given by :

$$j(\psi) := \mathcal{J}(u) = \frac{1}{2} \|v(u - u_d)\|_{\mathcal{X}}^2 \quad (5.1)$$

The state  $u(\psi)$  is related to the parameters  $\psi$  through an operator (that is likely to be non linear) that combines the partial differential equations along with the boundary conditions,

initial condition, etc. This operator is denoted as  $\mathcal{S}$  for the state problem. To be concise, one writes down

$$\mathcal{S}(u, \psi) = 0 \quad (5.2)$$

**Remark 13.** *If  $\mathbb{K}$  is  $\mathbb{C}$ , the forward model (5.2) is complex, and the  $\mathcal{X}$ -norm is a complex norm. Otherwise, if  $\mathbb{K}$  is  $\mathbb{R}$ , the forward model (5.2) is real, and the  $\mathcal{X}$ -norm is a real norm. In the following, the derivation is performed, with no loss of generality, assuming implicitly that  $\mathbb{K}$  is  $\mathbb{R}$ .*

In order to use efficient optimization algorithms for large scale inverse problems encountered in distributed parameter estimation, the cost function derivative or rather the cost function gradient has to be computed. To do so, the definition of the directional derivative of  $j(\psi)$  towards the direction  $\phi$  is used :

**Definition 3** (Directional derivative). *Let a point  $\psi$  and a direction  $\phi \in \mathcal{P}$ . The directional derivative of  $j$  at point  $\psi$  in the direction  $\phi$  is (Allaire et Craig, 2007) :*

$$j'(\psi; \phi) := \lim_{t \rightarrow 0^+} \frac{j(\psi + t\phi) - j(\psi)}{t} \quad (5.3)$$

The application of the directional derivative, eq. (5.3), on the cost function, eq. (5.1), leads to the equality :

$$j'(\psi; \phi) = (v(u - u_d), v'(u)u'(\psi; \phi))_{\mathcal{X}} \quad \forall \phi \in \mathcal{P} \quad (5.4)$$

and, because of the linearity of  $j'(\psi; \phi)$  with respect to  $\phi$  :

$$j'(\psi; \phi) = (\nabla j(\psi), \phi)_{\mathcal{X}} \quad \forall \phi \in \mathcal{P} \quad (5.5)$$

The inner product  $(\cdot, \cdot)_{\mathcal{X}}$  involved in the right-hand-side of eq. (5.5) will be also defined later on, when necessary. This is in this inner-product that regularization should be introduced.

The differentiation of the state problem in the direction  $\phi$  yields the directional derivative  $u'(\psi; \phi)$  that can further be injected into the inner product, eq (5.4). The state equation being



given by eq. (5.2), the derivative of  $S(u, \psi)$  toward the direction  $\phi$  is written as :

$$\mathcal{S}'_u(u, \psi)u'(\psi; \phi) + \mathcal{S}'_\psi(u, \psi)\phi = 0 \quad \forall \phi \in \mathcal{P} \quad (5.6)$$

where  $\mathcal{S}'_u$  is the tangent operator, i.e. the jacobian of  $\mathcal{S}$  with respect to  $u$  and  $\mathcal{S}'_\psi$  is the jacobian of  $\mathcal{S}$  with respect to  $\psi$ . Using this basic approach, one needs  $\dim \mathcal{P}$  integrations of eq. (5.6) to access the full gradient  $\nabla j(\psi) \in \mathbb{R}^{\dim \mathcal{P}}$  for a fully discretized problem. When  $\dim \mathcal{P}$  is large, which is likely to be the case when recovering space-distributed parameters, this approach becomes inefficient even though the computation of the linear tangent operator  $\mathcal{S}'_u(u, \psi)$  can be re-used for all directions  $\phi$  of the canonical basis of  $\psi$ .

Following Lions (Lions, 1968) and Céa (Céa, 1971), the adjoint-state method computes the whole cost function gradient through the integration of a single additional adjoint-state problem. A new variable (the adjoint-state variable  $u^*(\psi)$ ), is then introduced so that the cost function derivative also satisfies the relationship :

$$j'(\psi; \phi) = \left( \mathcal{S}'_\psi(u, \psi)\phi, u^* \right)_{\mathcal{Y}} \quad \forall \phi \in \mathcal{P} \quad (5.7)$$

where, again, the inner product  $(\cdot, \cdot)_{\mathcal{Y}}$  should be defined later on when necessary. Combining eq. (5.6) and eq. (5.7), the cost derivative rewrites :

$$j'(\psi; \phi) = - \left( \mathcal{S}'_u(u, \psi)u', u^* \right)_{\mathcal{Y}} \quad \forall \phi \in \mathcal{P} \quad (5.8)$$

The adjoint-state problem is then identified through eq. (5.4) and eq. (5.8), i.e. such that it satisfies the equality :

$$- \left( \mathcal{S}'_u(u, \psi)u', u^* \right)_{\mathcal{Y}} = \left( v(u - u_d), v'(u)u'(\psi; \phi) \right)_{\mathcal{X}} \quad \forall \phi \in \mathcal{P} \quad (5.9)$$

Next, if the adjoint problem, eq. (5.9), is satisfied (this means that the adjoint state  $u^*$  is accessible), then the cost function gradient is given by the inner product, eq. (5.7). In practice,

operators involved in eq. (5.9) are transposed. The inner product property  $(\mathcal{S}'_u(u, \psi)u', u^*)_{\mathcal{Y}} = (S^*(u, \psi)u^*, u')_{\mathcal{Y}}$  (with  $\mathcal{S}^*$  the transposed conjugate operator of  $\mathcal{S}'_u$ ) gives the adjoint equation concisely written as :

$$\mathcal{S}^*(u, \psi)u^* + v'(u)v(u - u_d) = 0 \quad \forall \phi \in \mathcal{P} \quad (5.10)$$

As reported by Protas (Protas, 2008), the gradient is identified in a given space for which the metric had been selected. To do so, the two following definitions are needed.

**Definition 4** ( $L^2(\Omega)$  Hilbert space).

$$L_2(\Omega) = \left\{ u : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} u(x)u(x) \, dx < \infty \right\} \quad (5.11)$$

*This functional space is associated to the inner product :*

$$(u, v)_{L_2(\Omega)} = \int_{\Omega} u(x)v(x) \, dx \quad (5.12)$$

**Definition 5** ( $H^1(\Omega)$  Hilbert space).

$$H^1(\Omega) := \{ \psi : \Omega \rightarrow \mathbb{R} \mid \psi \in L^2(\Omega) ; |\nabla \psi| \in L^2(\Omega) \} \quad (5.13)$$

*This functional space is associated to the inner product :*

$$(\psi, \phi)_{H^1(\Omega)} := (\psi, \phi)_{L^2(\Omega)} + (\nabla \psi, \nabla \phi)_{L^2(\Omega)} \quad (5.14)$$

Most often, the  $L^2(\Omega)$  real inner product for  $(\cdot, \cdot)_{\mathcal{X}}$  is used to extract the gradient. Combining eq. (5.5) and eq. (5.7) with the inner product of Definition 4 gives, after transposing within the right-hand-side inner product, the following result :

$$(\nabla j(\psi), \phi)_{L_2(\Omega)} = \left( \mathcal{S}'_{\psi}(u, \psi)u^*, \phi \right)_{\mathcal{Y}} \quad \forall \phi \in L^2(\Omega) \quad (5.15)$$

This way for writing down the cost function gradient is the ordinary one as first suggested by Lions (Lions, 1968) and C  a (C  a, 1971). However, according to Protas (Protas, 2008), it is suggested to use other inner product definitions due to the poor scaling of the corresponding discrete optimization problem. In some other studies, solving partial differential equations through optimization (see *e.g.* the articles of Mahavier (Mahavier, 1997), Danaila and Kazemi (Danaila et Kazemi, 2010) and Majid and Sial (Majid et Sial, 2010)), in image treatment (see *e.g.* Renka (Renka, 2009)), in boundary conditions reconstruction (see *e.g.* Nassiopoulou and Bourquin (Nassiopoulou et Bourquin, 2010) and Bourquin and Nassiopoulou (Bourquin et Nassiopoulou, 2011)) or in initial state reconstruction (Protas, 2008), the Sobolev space  $H^1(\Omega)$  is used for extracting the cost function gradient. This acts as a preconditioner modifying the direction of descent involved in the gradient-type optimization algorithm. It is shown, for instance in the study of Danaila and Kazemi (Danaila et Kazemi, 2010) that the cost function decrease when solving the PDEs problem can be much higher with the  $H^1$  inner product than with the ordinary  $L^2$  one.

In the present study, the choice of another inner product comes from the fact that we are faced with dealing with noisy measurements  $u_d$ . As a matter of fact, the measurements are noisy by nature. These measures are involved in the cost function definition, eq. (5.1), through a difference with the predictions. The noisy measurements are thus also involved in the forcing term  $v'(u)v(u - u_d)$ . The noise then propagates through the adjoint equation, eq. (5.10), and then to the cost function gradient, eq. (5.15).

In order to deal with the smoothing of the measurements noise that propagates in the adjoint system and then to the cost function gradient, the weighted Sobolev space is introduced.

**Definition 6** ( $H^{1(\vartheta^S)}(\Omega)$  Weighted Sobolev space inner product).

$$(\psi, \phi)_{H^{1(\vartheta^S)}(\Omega)} := (\psi, \phi)_{L^2(\Omega)} + \left( \vartheta^S \nabla \psi, \vartheta^S \nabla \phi \right)_{L^2(\Omega)} \quad (5.16)$$

**Remark 14.**  $\vartheta^S \in \mathbb{R}_+$  is the regularization parameter.

This strategy has been applied in bi-dimensional optical tomography applications based on the radiative transfer equation (Balima *et al.*, 2013) where it was shown how the diffusion

of the noise could attenuate fluctuations and thus give better reconstructions with a moderate parameter  $\vartheta^S$ .

The strategy that is developed here goes much further : it consists in filtering the cost function gradient where, and only where it is needed, i.e. where the high frequency fluctuations exist, at the vicinity of the sensors.

In applications considered here, i.e. in optical tomography, sensors are located on the boundaries, so that the effect of the noise appears on the boundaries before being diffused through the adjoint-state equation. Thus, the idea is to choose a filtering function whose value is high on the boundary and that continuously decreases within the medium. To do so, one uses the distance function definition :

**Definition 7** (Distance function). *Let  $\mathbf{r} \in \Omega$  and  $y \in \partial\Omega$ ;  $d(\mathbf{r}, \partial\Omega)$  is the Euclidean distance function to the set  $\partial\Omega$  if*

$$d(\mathbf{r}, \partial\Omega) = \min_{\mathbf{y} \in \partial\Omega} \|\mathbf{r} - \mathbf{y}\| \quad (5.17)$$

**Remark 15.** *As an example, the filtering function can be written as :*

$$\ell_v^2(\mathbf{r}) = \exp(-vd^2(\mathbf{r}, \partial\Omega)) \quad (5.18)$$

This makes possible the construction of the space-dependent weighted Sobolev space :

**Definition 8** ( $H^{1(v)}(\Omega)$  Space-dependent weighted Sobolev space inner product).

$$(\psi, \phi)_{H^{1(v)}(\Omega)} := (\psi, \phi)_{L^2(\Omega)} + \left( \vartheta_v^S(\mathbf{r}) \nabla \psi, \vartheta_v^S(\mathbf{r}) \nabla \phi \right)_{L^2(\Omega)} \quad (5.19)$$

**Remark 16.** *With the choice of the filtering function, eq. (5.18), the regularization parameter becomes  $v$ .*

Using the  $H^{1(v)}(\Omega)$  inner product for  $(\cdot, \cdot)_{\mathcal{H}}$ , and if  $\phi \in H^1(\Omega)$ , then exists a unique  $\nabla^{H^{1(v)}} j(\psi) \in H^{1(v)}(\Omega)$  such that

$$j'(\psi; \phi) = \left( \nabla^{H^{1(v)}} j(\psi), \phi \right)_{H^{1(v)}(\Omega)} \quad \forall \phi \in H^1(\Omega) \quad (5.20)$$

Taking into account that

$$\left( \nabla \nabla^{H^{1(v)}} j(\psi), \nabla \phi \right)_{L^2(\Omega)} = - \left( \Delta \nabla^{H^{1(v)}} j(\psi), \phi \right)_{L^2(\Omega)} + \left( \nabla \nabla^{H^{1(v)}} j(\psi) \cdot \mathbf{n}, \phi \right)_{L^2(\partial\Omega)} \quad (5.21)$$

one obtains by identifying terms :

$$\left( \left( 1 - \vartheta^{S_v^2} \Delta \right) \nabla^{H^{1(v)}} j(\psi), \phi \right)_{L^2(\Omega)} = \left( \mathcal{J}'_\psi(u, \psi) u^*, \phi \right)_{\mathcal{Y}} \quad \forall \phi \in H^1(\Omega) \quad (5.22)$$

with a null flux boundary condition, i.e.  $\nabla \nabla^{H^{1(v)}} j(\psi) \cdot \mathbf{n} = 0$  on  $\partial\Omega$ . At this stage, the application being defined (measurements, state model, etc.), the  $\mathcal{Y}$ -norm is given accordingly, and the cost function gradient can be extracted.

**Remark 17.** *Due to the inclusion  $H^{1(v)}(\Omega) \subset L^2(\Omega)$ , the space-dependent weighted Sobolev gradient  $\nabla^{H^{1(v)}} j(\psi)$  is indeed an acceptable direction of descent for the optimization problem that consists in minimizing eq. (5.1).*

### 5.3 Application on 2D RTE-based optical tomography

#### 5.3.1 Inverse problem statement

The inverse problem consists mathematically in minimizing a cost function which quantitatively measures the discrepancy between radiative intensity measurements  $\check{L}$  and related predictions  $L$ . For optical tomography applications based on the radiative transfer equation, this cost function, eq. (5.23), mapping from  $\mathcal{P} = L^\infty(\Omega) \times L^\infty(\Omega)$  to  $\mathbb{R}$  uses, according to the schematic description of the problem (Figure 5.1), the specific Euclidean norm :

$$j(\kappa, \sigma_s) := \mathcal{J}(L) = \frac{1}{2} \left\| \frac{\mathbf{s} \cdot \mathbf{n}}{\check{L}} (L - \check{L}) \right\|_{\mathcal{X}}^2 \quad \text{with} \quad \|u\|_{\mathcal{X}} \triangleq \left( \sum_{k=1}^K \sum_{d=1}^D \sum_{r=1}^R u_k(\mathbf{r}_d, \mathbf{s}_r)^2 \right)^{\frac{1}{2}} \quad (5.23)$$

where the index  $k$  refers to the source (test) number, the index  $d$  refers to the detection node and the index  $r$  refers to the direction. For a given source configuration, say the  $k^{\text{th}}$  one, the prediction  $L_k(\mathbf{r}_d, \mathbf{s}_r)$  at the sensor  $d$  and direction  $r$  is given through the solution of the forward

steady-state radiative transfer problem formulated as :

$$\left\{ \begin{array}{l} \text{Find } L_k(\mathbf{r}, \mathbf{s}) \text{ such that } \forall \mathbf{s} \in \mathcal{S}^{n_\Omega-1}, \forall k = 1, \dots, K : \\ (\mathbf{s} \cdot \nabla + \kappa + \sigma_s) L_k(\mathbf{r}, \mathbf{s}) = \sigma_s \int_{\mathcal{S}^{n_\Omega-1}} L_k(\mathbf{r}, \mathbf{s}') \Psi(\mathbf{s}, \mathbf{s}') d\mathbf{s}', \forall \mathbf{r} \in \Omega \\ L_k(\mathbf{r}, \mathbf{s}) = L_k^0(\mathbf{r}) \times \mathbb{1}_{[\mathbf{r} \in \partial\Omega_k^0]} \times \mathbb{1}_{[\mathbf{s} = \mathbf{s}_k^0]}, \forall \mathbf{r} \in \partial\Omega \end{array} \right. \quad (5.24)$$

where  $\mathbf{s}$  is the considered direction of propagation,  $\Psi(\mathbf{s}, \mathbf{s}')$  is the phase function representing the probability that a photon arriving from the direction  $\mathbf{s}'$  is scattered to the direction  $\mathbf{s}$ ,  $\kappa$  and  $\sigma_s$  are the absorption and scattering space-dependent functions, respectively, and  $\mathbb{1}_{[\cdot]}$  denotes the indicator function.

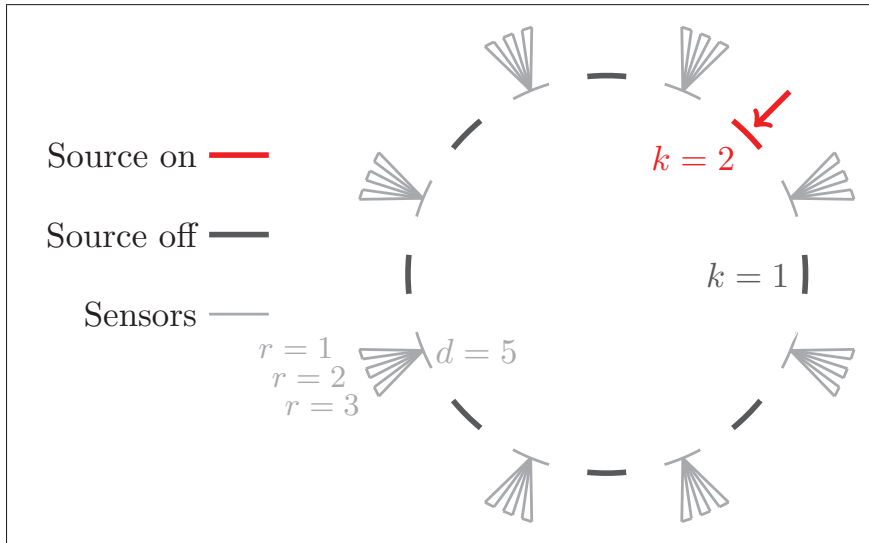


FIG. 5.1 Schematic description of the experiment. Only one source is on at once, while others are off. For each source configuration number  $k = 1, \dots, K$ , the emerging intensity is measured on all sensors  $d = 1, \dots, D$  in different directions  $r = 1, \dots, R$ .

### 5.3.2 Cost function gradient derivation

The computation of the cost function gradient follows the generic methodology detailed in the previous section. The cost function directional derivative is given by :

$$j'(\boldsymbol{\kappa}, \boldsymbol{\sigma}_s; \phi) = (\nabla j(\boldsymbol{\kappa}, \boldsymbol{\sigma}_s), \phi)_{\mathcal{X}} = \left( \left( \frac{\mathbf{s} \cdot \mathbf{n}}{\check{L}} \right)^2 (L - \check{L}), L' \right)_{\mathcal{X}} \quad \forall \phi \in \mathcal{P} \quad (5.25)$$

In order to derive the cost gradient for this specific application,  $\mathcal{Y}$ -based inner products are defined :

**Definition 9.**

$$(u, v)_{\mathcal{Y}} \triangleq \sum_k \int_{\mathcal{S}^{n_{\Omega}-1}} \int_{\Omega} uv \, d\mathbf{r} \, d\mathbf{s} \quad (5.26)$$

$$(u, v)_{\partial \mathcal{Y}} \triangleq \sum_k \int_{\mathcal{S}^{n_{\Omega}-1}} \int_{\partial \Omega} uv \, d\mathbf{r} \, d\mathbf{s} \quad (5.27)$$

The  $\mathcal{Y}$ -inner product satisfies properties 1 and 2, both being useful for the cost function gradient derivation.

**Property 1.**

$$(\mathbf{s} \cdot \nabla u, v)_{\mathcal{Y}} = -(\mathbf{s} \cdot \nabla v, u)_{\mathcal{Y}} + (\mathbf{s} \cdot \mathbf{n}v, u)_{\partial \mathcal{Y}} \quad (5.28)$$

**Property 2.**

$$\begin{aligned} \left( \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} u(\mathbf{r}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') \, d\mathbf{s}', v(\mathbf{r}, \mathbf{s}) \right)_{\mathcal{Y}} &= \\ \sum_k \int_{\mathcal{S}^{n_{\Omega}-1}} \int_{\Omega} \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} u(\mathbf{r}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') \, d\mathbf{s}' v(\mathbf{r}, \mathbf{s}) \, d\mathbf{r} \, d\mathbf{s} &= \\ \sum_k \int_{\Omega} \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} \int_{\mathcal{S}^{n_{\Omega}-1}} v(\mathbf{r}, \mathbf{s}) u(\mathbf{r}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') \, d\mathbf{s}' \, d\mathbf{s} \, d\mathbf{r} &= \\ \sum_k \int_{\Omega} \int_{\mathcal{S}^{n_{\Omega}-1}} \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} v(\mathbf{r}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\mathbf{s} u(\mathbf{r}, \mathbf{s}') \, d\mathbf{r} \, d\mathbf{s}' &= \\ \left( \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} v(\mathbf{r}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\mathbf{s}', u(\mathbf{r}, \mathbf{s}') \right)_{\mathcal{Y}} \end{aligned}$$

The generic equation for the cost gradient, eq. (5.15), applied on both space-distributed parameters  $\kappa$  and  $\sigma_s$  gives, for direction  $\phi = (\eta_\kappa; 0) \in \mathcal{P}$  :

$$(\nabla j(\kappa, \sigma_s), \phi)_{\mathcal{X}} = (L(\mathbf{r}, \mathbf{s})L^*(\mathbf{r}, \mathbf{s}), \phi)_{\mathcal{Y}} \quad (5.29)$$

and for direction  $\phi = (0; \eta_{\sigma_s}) \in \mathcal{P}$  :

$$(\nabla j(\kappa, \sigma_s), \phi)_{\mathcal{X}} = \left( L(\mathbf{r}, \mathbf{s})L^*(\mathbf{r}, \mathbf{s}) - \int_{\mathcal{S}^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s})\phi(s, s')ds' L^*(\mathbf{r}, \mathbf{s}), \phi \right)_{\mathcal{Y}} \quad (5.30)$$

On one hand, eq. (5.29) is rewritten such that

$$(\nabla j(\kappa, \sigma_s), \phi)_{\mathcal{X}} = (\phi L(\mathbf{r}, \mathbf{s}), L^*(\mathbf{r}, \mathbf{s}))_{\mathcal{Y}} \quad (5.31)$$

and, on the other hand, property 2 enables to rewrite eq. (5.30) such that :

$$(\nabla j(\kappa, \sigma_s), \phi)_{\mathcal{X}} = \left( \phi L(\mathbf{r}, \mathbf{s}) - \phi \int_{\mathcal{S}^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s})\phi(s, s')ds', L^*(\mathbf{r}, \mathbf{s}) \right)_{\mathcal{Y}} \quad (5.32)$$

Besides, the differentiation of the forward model enables to rewrite previous equations too, for  $\phi$  being either  $(\eta_\kappa; 0)$  or  $(0; \eta_{\sigma_s})$  :

$$\begin{aligned} (\nabla j(\kappa, \sigma_s), \phi)_{\mathcal{X}} = & -((\mathbf{s} \cdot \nabla + \kappa + \sigma_s)L'(\mathbf{r}, \mathbf{s}), L^*)_{\mathcal{Y}} \\ & + \left( \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} L'(\mathbf{r}, \mathbf{s})\phi(\mathbf{s}, \mathbf{s}') d\mathbf{s}', L^* \right)_{\mathcal{Y}} \end{aligned} \quad (5.33)$$

Properties 1 and 2 enables to rewrite eq. (5.33) such that :

$$\begin{aligned} (\nabla j(\kappa, \sigma_s), \phi)_{\mathcal{X}} = & ((\mathbf{s} \cdot \nabla - \kappa - \sigma_s)L^*(\mathbf{r}, \mathbf{s}), L')_{\mathcal{Y}} - (\mathbf{s} \cdot \mathbf{n}L^*, L')_{\partial \mathcal{Y}} \\ & + \left( \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} L^*(\mathbf{r}, \mathbf{s})\phi(\mathbf{s}, \mathbf{s}') d\mathbf{s}', L' \right)_{\mathcal{Y}} \end{aligned} \quad (5.34)$$



The adjoint problem is eventually obtained identifying eq. (5.25) with eq. (5.34) :

$$\left| \begin{array}{l} \text{Find } L_k(\mathbf{r}, \mathbf{s}) \text{ such that } \forall \mathbf{s} \in \mathcal{S}^{n_{\Omega}-1}, \forall k = 1, \dots, k : \\ (-\mathbf{s} \cdot \nabla + \kappa + \sigma_s) L_k^*(\mathbf{r}, \mathbf{s}) = \sigma_s \int_{\mathcal{S}^{n_{\Omega}-1}} L_k^*(\mathbf{r}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') d\mathbf{s}', \forall \mathbf{r} \in \Omega \\ -\mathbf{s} \cdot \mathbf{n} L_k^*(\mathbf{r}, \mathbf{s}) = \left( \frac{\mathbf{s} \cdot \mathbf{n}}{L} \right)^2 (L - \check{L}) \times \mathbb{1}_{[\mathbf{r} \in \partial\Omega_d]}(\mathbf{r}), \forall \mathbf{r} \in \partial\Omega \end{array} \right. \quad (5.35)$$

where  $\partial\Omega_d$  denotes the set of the pointwise sensors.

### 5.3.3 Numerical results

The optimization relies on the L-BFGS algorithm described by Liu and Nocedal (Liu et Nocedal, 1989) coupled with a quadratic interpolation line-search (Antonioni et Lu, 2007). Generally speaking, this latter has proved to be particularly effective these recent years in the area of inversion (see (Wang *et al.*, 2009; Mohebbi et Sellier, 2014; Jin *et al.*, 2015) to name but a few) due to its effectiveness to minimize nonlinear problems (Liu, 2014).

The two-dimensional domain is a square of 2 cm length. Four collimated sources are located at the center of each side, while sensors are placed around sources towards twenty directions. The target properties to be reconstructed,  $\kappa^t$  and  $\sigma_s^t$ , are defined by the background for which  $\kappa = 0.25 \text{ cm}^{-1}$  and  $\sigma_s = 20 \text{ cm}^{-1}$  and two square inclusions, the former for which  $\kappa = 0.15 \text{ cm}^{-1}$  and  $\sigma_s = 10 \text{ cm}^{-1}$  and the latter for which  $\kappa = 0.35 \text{ cm}^{-1}$  and  $\sigma_s = 30 \text{ cm}^{-1}$ . Both the forward and the adjoint models are solved using  $\mathcal{P}_1$  discontinuous Galerkin finite elements for which the mesh convergence has been validated. Three distinct meshes have been created, one in order to build the synthetic data before adding random noise following Klose (Klose, 2001) (3721 nodes), one in order project the forward and adjoint state variables (1681 nodes), and one associated to the property functions with the aim of implicitly regularize the inverse problem (441 nodes). Seeking to gauge the accuracy of the reconstructions, the deviation factor  $e_1 \in [0, +\infty[$  and the correlation factor  $e_2 \in [-1, +1]$  are used (Flannery *et al.*, 1992). The former factor measures the deviation of the reconstructed image from the target image : a small value indicates a reconstructed image with high accuracy. The latter factor measures

the linear correlation between the target and the reconstructed image : a large value shows a high detectability in the reconstructed image and indicates a reconstructed image with high accuracy.

Numerical tests were performed for different levels of noise. First, a small noise of 20 dB is added to the synthetic data (this corresponds to 1 % of noise). Then, a moderate noise of 15 dB is added to the data (approximately 3 % of noise). Finally, the highest level of noise of 10 dB is considered (10 % of noise).

Figure 5.2 presents the evolution of dimensionless cost function values along with iterations. Note that the main used stopping criterion has been based on the cost function stabilization, with a critical value chosen equal to  $10^{-2}$ . As expected, it is seen from this figure that the higher the level of noise, the higher the cost function value at stabilization. Moreover, the convergence to the solution is slightly faster with Sobolev gradients than with ordinary ones, even though, far at the end, the value of the cost function is the lowest for the ordinary gradient.

Figures 5.3, 5.4 and 5.5 present the evolutions, with respect to iterations, of deviation and correlation factors, respectively to both parameters  $\kappa$  and  $\sigma_s$ , for the three levels of noise. These results show that when compared to ordinary gradients, the use of space-dependent Sobolev gradients decreases deviation factors and increases correlation factors. The optimum regularization parameter,  $\nu$ , is about 100 in this particular case : when  $\nu$  is too small, there is too much gradient diffusion while, when  $\nu$  is too high, this yields the ordinary  $L^2(\Omega)$  gradient.

Figures 5.6 and 5.7 corroborate these conclusions. The Sobolev gradients filter the high frequency fluctuations and avoid the propagation of the noise to the cost function gradient and then to the reconstructions. It is observed that the regularization parameter  $\nu = 100$  gives the best reconstructions as regards smoothness and precision of reconstructions.

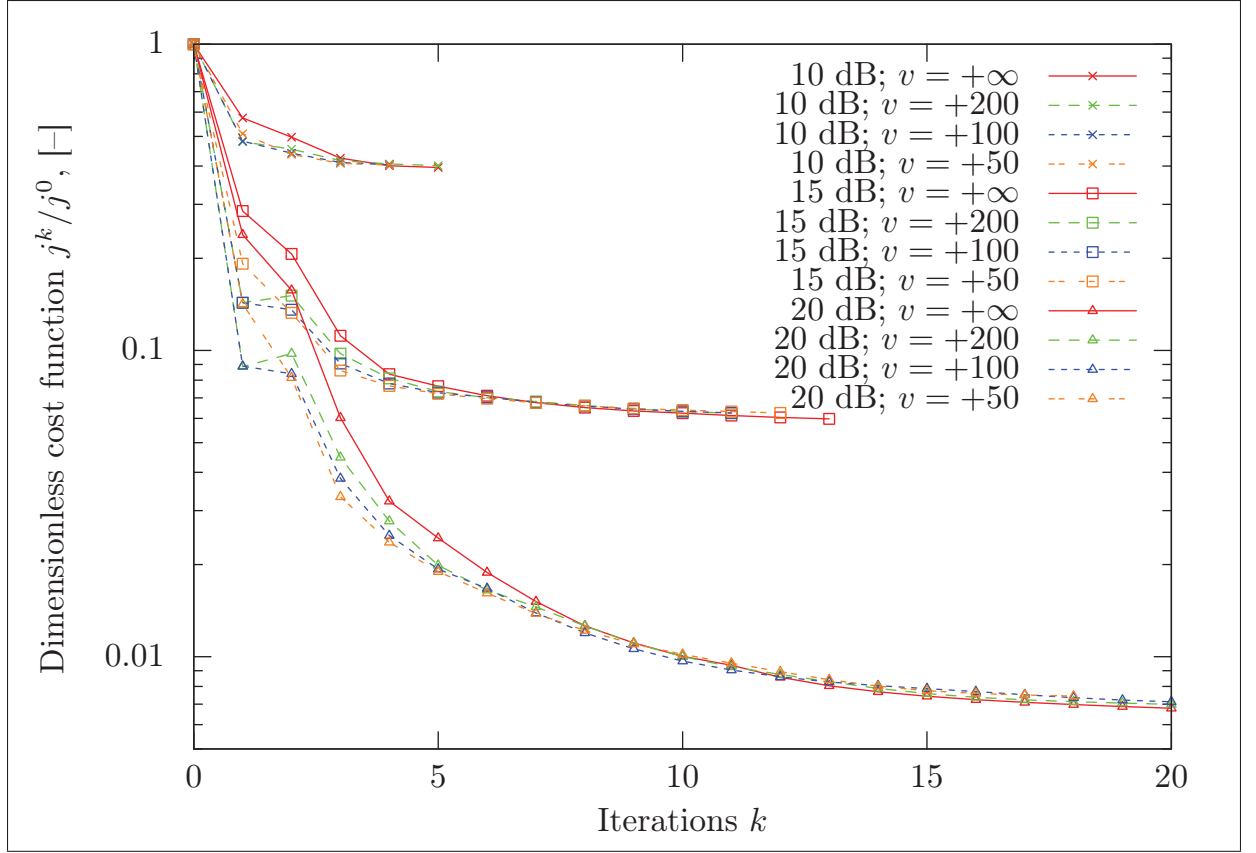


FIG. 5.2 Evolution of the cost function with respect to iterations, for  $v = +\infty, 200, 100$  and  $50$  and a 20 dB, 15 dB and 10 dB noise.

The algorithms stops at iteration  $k$  when  $\frac{|j^k - j^{k-1}|}{j^{k-1}} < 10^{-2}$ .

## 5.4 Application on 3D DA-based optical tomography

### 5.4.1 Inverse problem statement

The radiative transfer equation (RTE) provides an equation of light propagation valid in most of participating media as long as the independent scattering regime is fulfilled (Sacadura, 2011). For specific applications, approximations of the RTE could be applied. Solutions of the RTE being computationally expensive, especially if the third dimension has to be considered, approximated models of the RTE should be employed when valid. Among the different models approaching the RTE, the diffuse approximation (DA) is the privileged one. The latter, which describes the photon density field  $\varphi(\mathbf{r}) = \int_{\mathcal{S}^{n_{\Omega}-1}} L(\mathbf{r}, \mathbf{s}) \, d\mathbf{s}$  inside the medium, is given in the

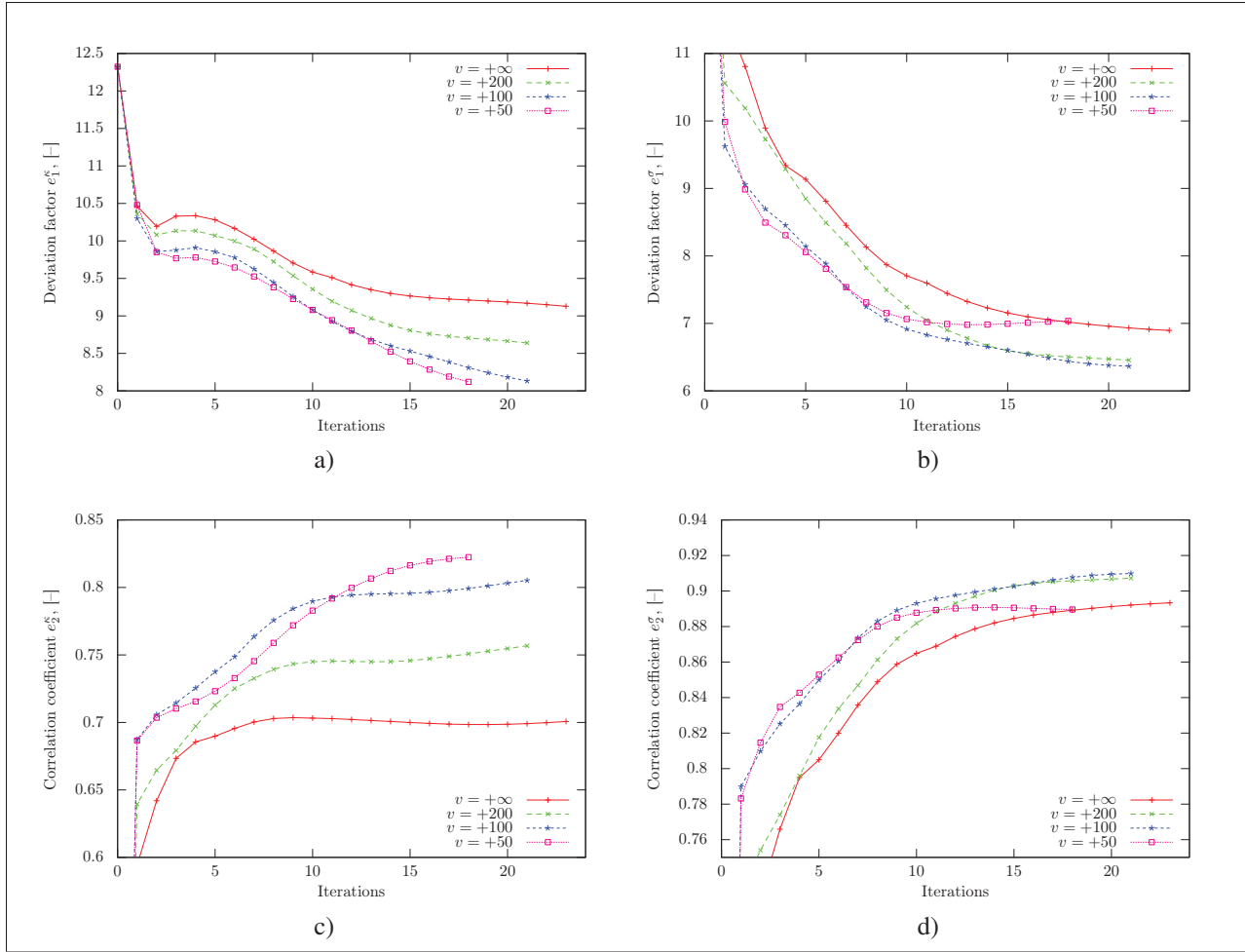


FIG. 5.3 20 dB. Evolution of errors  $e_1$  (top) and  $e_2$  (bottom) for  $\kappa$  (left) and  $\sigma_s$  (right), with respect to iterations, for  $v = +\infty, 200, 100$  and  $50$ . The algorithms stops at iteration  $k$  when

$$\frac{|j^k - j^{k-1}|}{j^{k-1}} < 10^{-2}.$$

frequency domain by (Arridge, 1999; Kaipio et Somersalo, 2005) :

$$\begin{cases} -\nabla \cdot [D(\mathbf{r}) \nabla \hat{\varphi}_k(\mathbf{r}, f_0)] + \left[ \kappa(\mathbf{r}) + \frac{2\pi i f_0}{c} \right] \hat{\varphi}_k(\mathbf{r}, f_0) = \hat{X}_k(\mathbf{r}, f_0), \forall \mathbf{r} \in \Omega \\ 2\gamma \hat{\varphi}_k(\mathbf{r}, f_0) + AD(\mathbf{r}) \nabla \hat{\varphi}_k(\mathbf{r}, f_0) \cdot \mathbf{n} = 2\hat{x}_k(\mathbf{r}, f_0), \forall \mathbf{r} \in \partial\Omega \end{cases} \quad (5.36)$$

Here,  $\hat{X}_k(\mathbf{r}, f_0)$  is the Fourier transform at the frequency  $f_0$  of the internal source of radiation,  $X_k$  [ $\text{W cm}^{-3}$ ];  $\hat{x}_k(\mathbf{r}, f_0)$  is the Fourier transform at the frequency  $f_0$  of the temporal signal associated with the  $k^{\text{th}}$  diffuse source,  $x_k$  [ $\text{W cm}^{-2}$ ];  $\hat{\varphi}_k : (\Omega \times \mathbb{C}) \mapsto \mathbb{R}$  is the Fourier transform

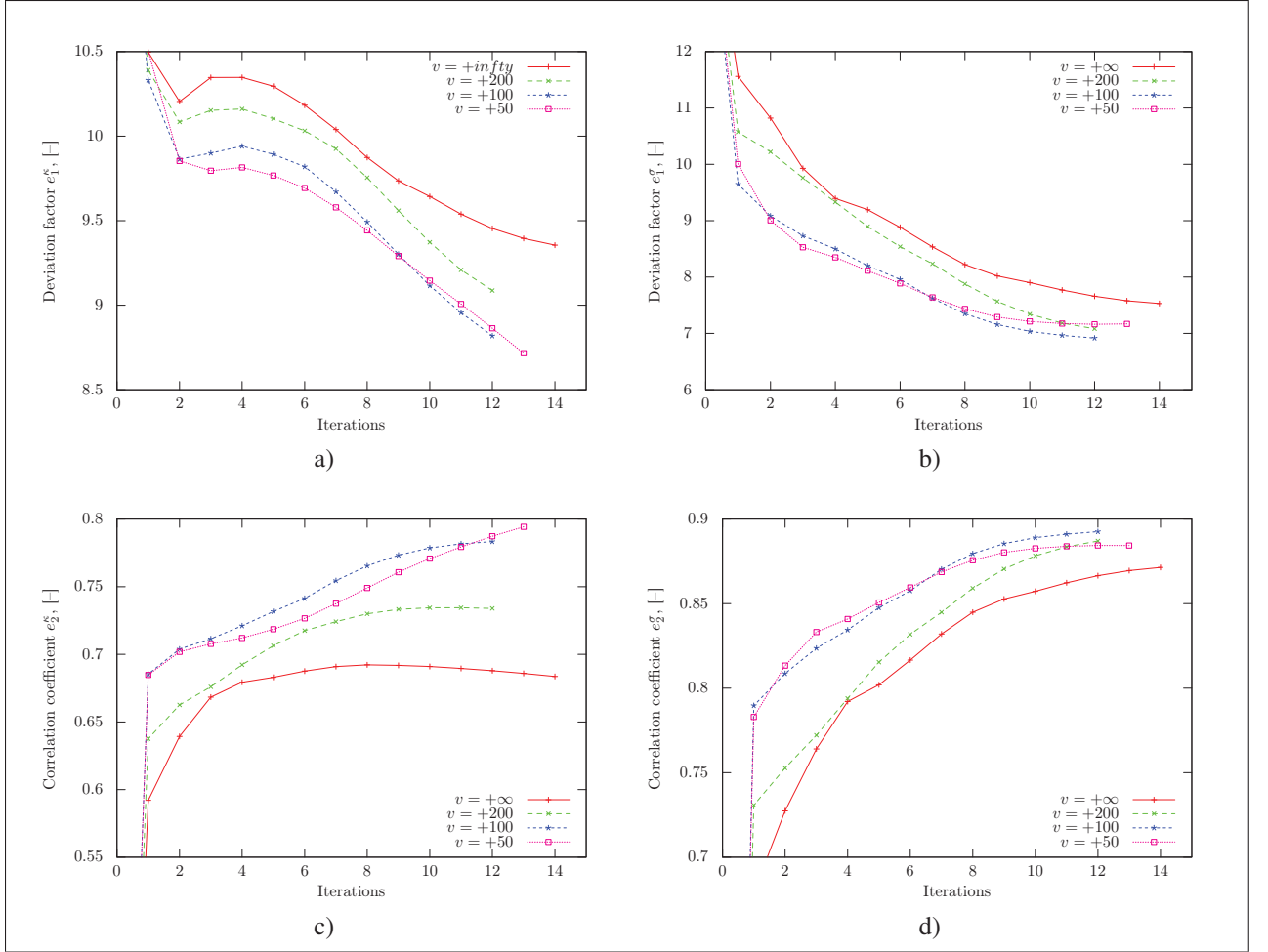


FIG. 5.4 15 dB. Evolution of errors  $e_1$  (top) and  $e_2$  (bottom) for  $\kappa$  (left) and  $\sigma_s$  (right), with respect to iterations, for  $v = +\infty, 200, 100$  and  $50$ . The algorithm stops at iteration  $k$  when

$$\frac{|j^k - j^{k-1}|}{j^{k-1}} < 10^{-2}.$$

of the  $k^{\text{th}}$  photon density field,  $\phi_k : (\Omega \times \mathbb{R}^+) \mapsto \mathbb{R} [\text{W cm}^{-2}]$ ;  $D(\mathbf{r}) = (n_\Omega [\kappa(\mathbf{r}) + \sigma(\mathbf{r})])^{-1}$  is the macroscopic scattering coefficient  $[\text{cm}]$ ;  $n_\Omega$  ( $= 2, 3$ ) is the dimension of  $\Omega$ ;  $\kappa$  and  $\sigma = (1 - g)\sigma_s$  are the absorption and reduced scattering coefficients ( $\sigma_s$  is the scattering coefficient as previously defined)  $[\text{cm}^{-1}]$ ;  $g$  is the asymmetry factor which equals to the average cosine of the scattering angle;  $c$  is the speed of light in  $\Omega$  (constant)  $[\text{cm s}^{-1}]$ ;  $\mathbf{n}$  is the unit normal vector to the boundary of  $\Omega$ ;  $\gamma$  is a constant parameter depending on  $n_\Omega$  ( $1/\pi$  for  $n_\Omega = 2$ ,  $1/4$  for  $n_\Omega = 3$ ); and  $A$  is a parameter which characterizes the reflection on the frontier  $\partial\Omega$  that

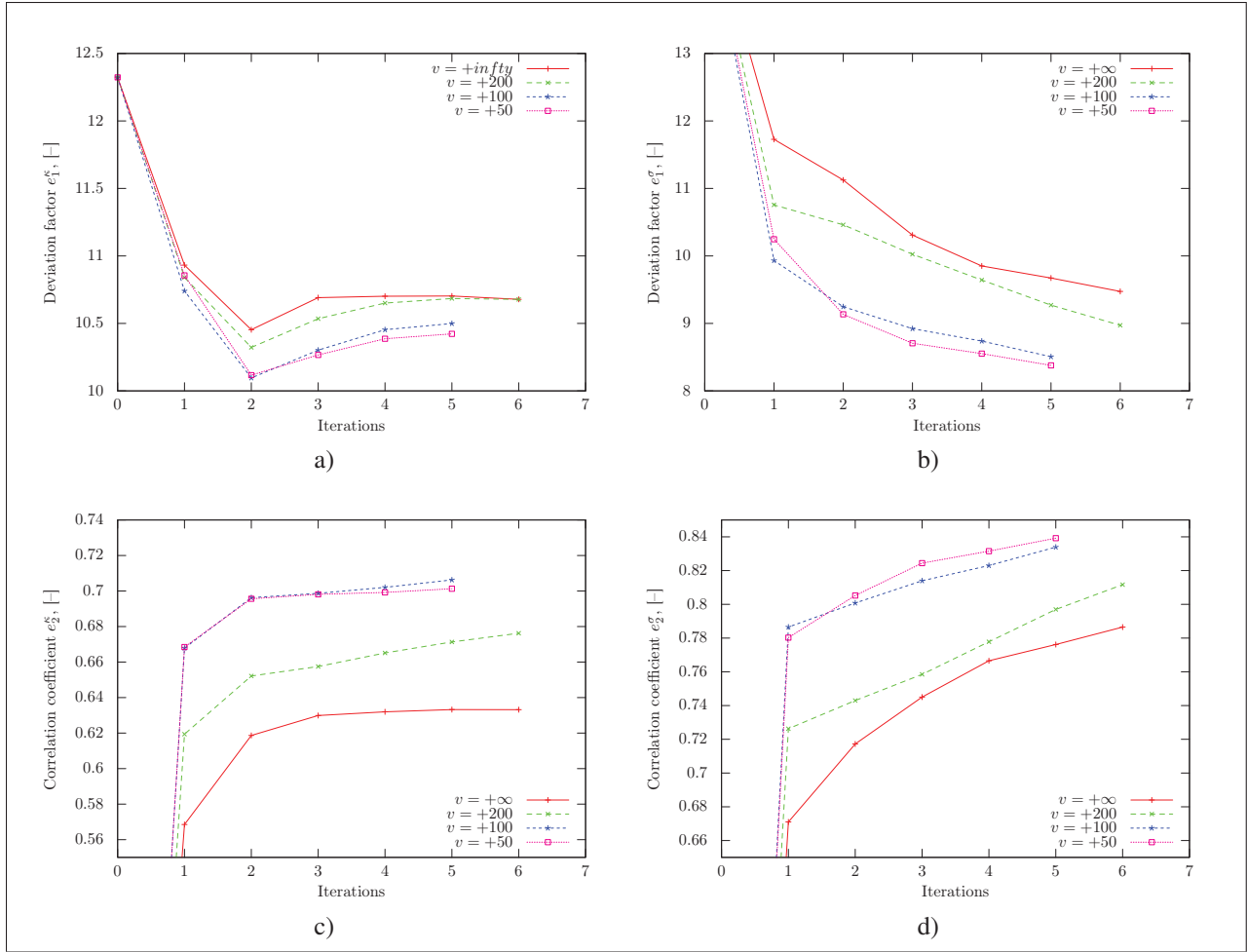


FIG. 5.5 10 dB. Evolution of errors  $e_1$  (top) and  $e_2$  (bottom) for  $\kappa$  (left) and  $\sigma_s$  (right), with respect to iterations, for  $v = +\infty, 200, 100$  and  $50$ . The algorithms stops at iteration  $k$  when

$$\frac{|j^k - j^{k-1}|}{j^{k-1}} < 10^{-2}.$$

can be derived from the Fresnel laws if specular reflection is considered (Keijzer *et al.*, 1988; Haskell *et al.*, 1994).

Similarly to section 5.3.3, the inverse problem consists in solving the optimization problem :

$$\text{Find } (\kappa, \sigma) \in \mathcal{P} = L^\infty(\Omega) \times L^\infty(\Omega) \text{ such that } (\kappa, \sigma) = \arg \min_{(\kappa, \sigma) \in \mathcal{P}} j(\kappa, \sigma) \quad (5.37)$$

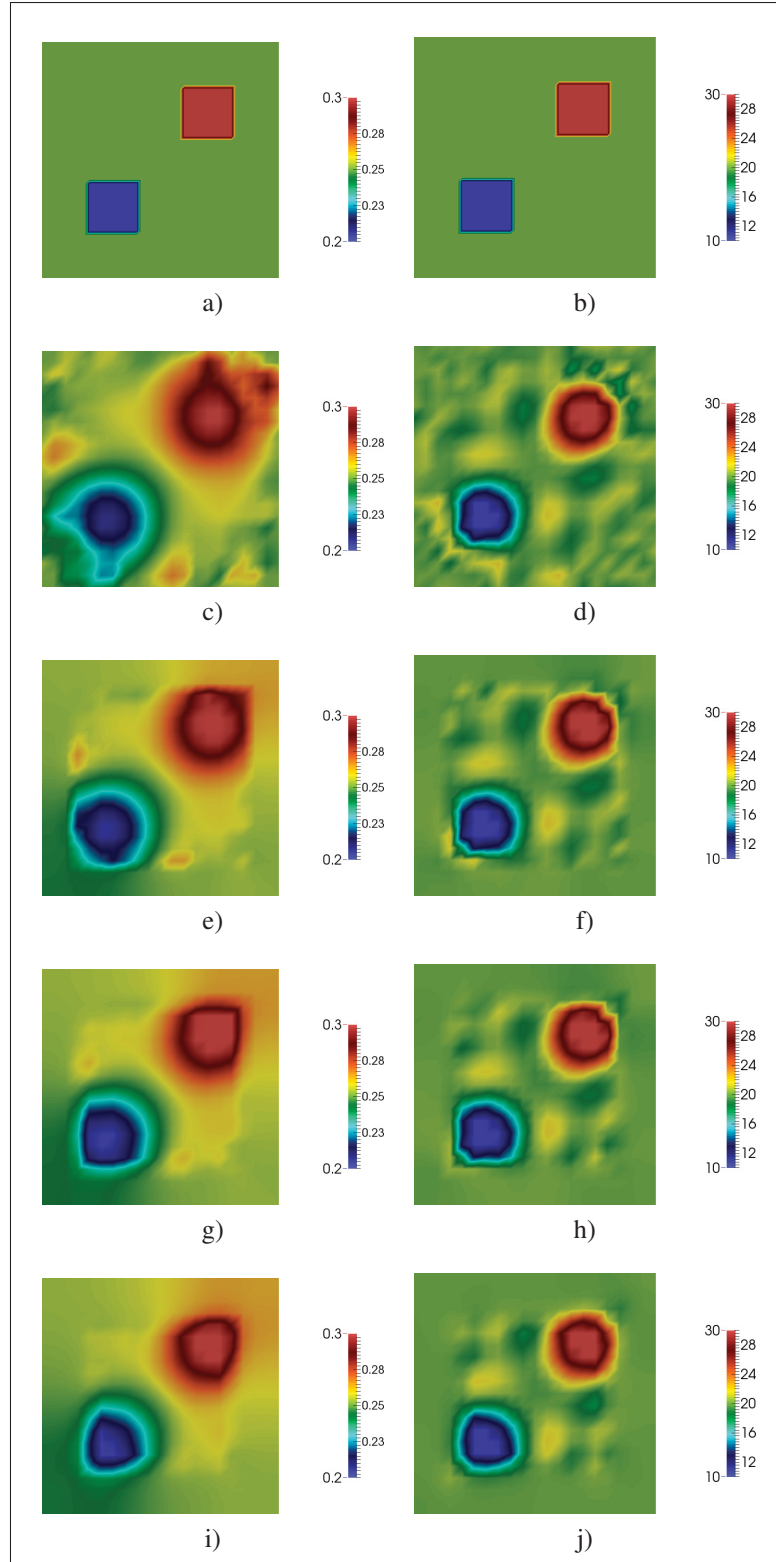


FIG. 5.6 20 dB. Reconstructions obtained for  $\kappa$  (left) and  $\sigma_s$  (right). 1st row : targets, then :  $\nu = +\infty$ , 200, 100 and 50, respectively.

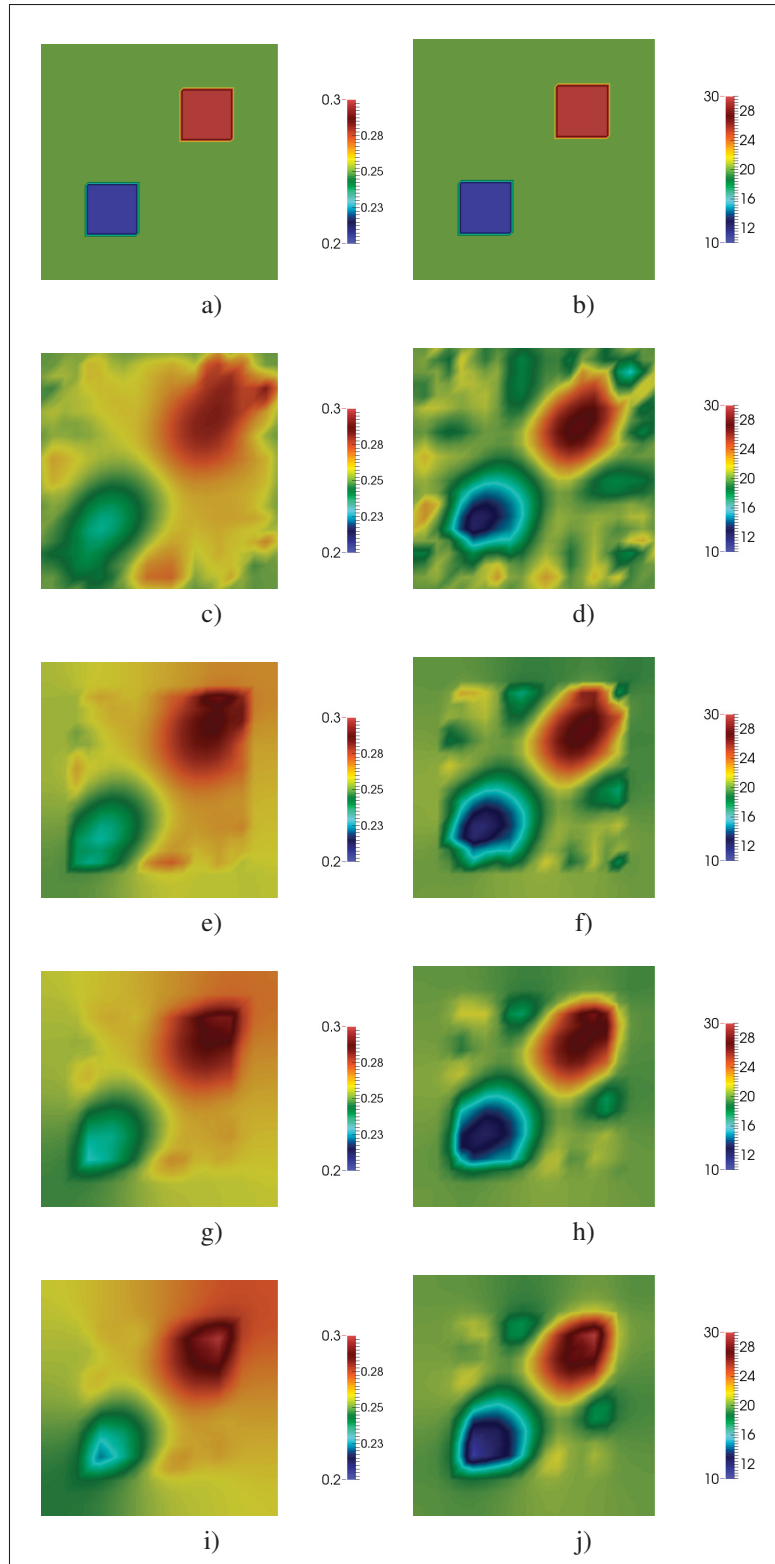


FIG. 5.7 10 dB. Reconstructions obtained for  $\kappa$  (left) and  $\sigma_s$  (right). 1st row : targets, then :  $\nu = +\infty$ , 200, 100 and 50, respectively.



where

$$j(\kappa, \sigma) := \mathcal{J}(\hat{\phi}) = \frac{1}{2} \left\| \frac{\hat{\phi} - \tilde{\phi}}{\hat{\phi}} \right\|_{\mathcal{X}}^2 \quad \text{with} \quad \|u\|_{\mathcal{X}} \triangleq \left( \sum_{k=1}^K \sum_{d=1}^D |u_k(\mathbf{r}_d)|^2 \right)^{\frac{1}{2}} \quad (5.38)$$

and where the index  $k$  refers to the source number and the index  $d$  refers to the detection node. It should be noted that, under the assumption of the DA, the measurable quantity in optical tomography,  $\Psi_{\text{out}}(\mathbf{r}) = \int_{\{\mathbf{s} \cdot \mathbf{n} > 0\}} (1 - R(\mathbf{r}, \mathbf{s})) L(\mathbf{r}, \mathbf{s}) \mathbf{s} \cdot \mathbf{n} d\mathbf{s}$ , is related to the photon density through the equality  $\Psi_{\text{out}}(\mathbf{r}) = 2\gamma\phi(\mathbf{r})$  (Kaipio et Somersalo, 2005). Here, the directional function  $R(\mathbf{r}, \mathbf{s})$  is the Fresnel reflection factor.

In the remainder of the section, the following values of the source terms are considered :  $\hat{X}_k = 0 \text{ W cm}^{-2}$  and  $\hat{x}_k = 0.1 \text{ W cm}^{-2}$  for all  $k = 1, \dots, K$  (which corresponds to a temporal signal of the form  $\cos(2\pi f_0 t)$ ).

#### 5.4.2 Cost function gradients derivation

The cost function gradients with respect to  $\kappa$  and  $\sigma$  have been computed in Dubot and al. (Dubot *et al.*, 2015b) with the help of a continuous Lagrangian formulation. In terms of eq. (5.5), it can be shown that the directional derivatives of the cost function are equal to :

$$(\nabla j(\psi), \phi)_{\mathcal{X}} = \sum_{k=1}^K \left( \text{Re} \left( \hat{\phi}_k \tilde{\phi}_k^* - n_{\Omega} D^2 \nabla \hat{\phi}_k \cdot \nabla \tilde{\phi}_k^* \right), \phi \right)_{\mathcal{X}} \quad \forall \phi = (\eta_{\kappa}; 0) \in \mathcal{P} \quad (5.39)$$

$$(\nabla j(\psi), \phi)_{\mathcal{X}} = \sum_{k=1}^K \left( \text{Re} \left( -n_{\Omega} D^2 \nabla \hat{\phi}_k \cdot \nabla \tilde{\phi}_k^* \right), \phi \right)_{\mathcal{X}} \quad \forall \phi = (0; \eta_{\sigma}) \in \mathcal{P} \quad (5.40)$$

where  $\text{Re}$  is the real part operator and the adjoint variable  $\tilde{\phi}_k^*$  involved in eq. (5.15) is the solution to the adjoint problem :

$$\left| \begin{array}{ll} -\nabla \cdot [D \nabla \hat{\phi}_k^*] + \left[ \kappa - \frac{2\pi i f_0}{c} \right] \hat{\phi}_k^* &= 0 \quad \forall \mathbf{r} \in \Omega \\ \frac{2\gamma}{A} \hat{\phi}_k^* + D \nabla \hat{\phi}_k^* \cdot \mathbf{n} &= -\frac{1}{|\hat{\phi}_k|^2} \left( \hat{\phi}_k - \tilde{\phi}_k \right) \mathbb{1}_{[\mathbf{r} \in \partial \Omega_d]}(\mathbf{r}) \quad \forall \mathbf{r} \in \partial \Omega \end{array} \right. \quad (5.41)$$

Therefore, if the inner product  $\mathcal{L} = L^2(\Omega)$  is used to extract the cost function gradient in eq. (5.5), the expressions for the continuous cost function gradients are such that :

$$\begin{cases} \nabla_{\kappa} j = \sum_{k=1}^K \operatorname{Re} \left( \widehat{\phi}_k \bar{\phi}_k^* - n_{\Omega} D^2 \nabla \widehat{\phi}_k \cdot \nabla \bar{\phi}_k^* \right) \\ \nabla_{\sigma} j = \sum_{k=1}^K \operatorname{Re} \left( -n_{\Omega} D^2 \nabla \widehat{\phi}_k \cdot \nabla \bar{\phi}_k^* \right) \end{cases} \quad (5.42)$$

Note that an adimensionalisation of the control variables by *a priori* radiative properties is used in the inversion process (Dubot *et al.*, 2015b).

### 5.4.3 Numerical results

The same algorithm as employed in section 5.3 is used to reconstruct the target properties from initial estimates equal to values of those of the background. The domain considered,  $\Omega$ , is a square of 5 cm length ( $\Omega = [0; 5]^2 \text{ cm}^2$ ). The target properties to be reconstructed,  $\kappa_t$  and  $\sigma_t'$ , are defined by :

$$\begin{cases} \kappa_t(\mathbf{r}) = 0.1 + 0.05 \times \mathbb{1}_{[\mathbf{r} \in \Omega_1]}(\mathbf{r}) + 0.1 \times \mathbb{1}_{[\mathbf{r} \in \Omega_2]}(\mathbf{r}) \text{ cm}^{-1} \\ \sigma_t'(\mathbf{r}) = 10 + 5 \times \mathbb{1}_{[\mathbf{r} \in \Omega_1]}(\mathbf{r}) + 10 \times \mathbb{1}_{[\mathbf{r} \in \Omega_2]}(\mathbf{r}) \text{ cm}^{-1} \end{cases} \quad (5.43)$$

where  $\Omega_1 = [0.75; 1.75]^2 \text{ cm}^2$  and  $\Omega_2 = [3.25; 4.25]^2 \text{ cm}^2$ .

The stopping criterion are based on the cost function stabilization, with a critical value chosen equal to  $10^{-3}$  and a total number of iterations equals 100. Physical properties involved in the forward model, eq. (5.36), are fixed to :  $f_0 = 100 \text{ MHz}$ ,  $c = (3 \times 10^{10} / 1.4) \text{ cm s}^{-1}$  and  $\gamma = 0.25$ . Synthetic data are considered for these numerical tests. These data,  $\widehat{\phi}$ , which representing pseudo-experimental measurements, are built on a finer mesh (132651 nodes) than the one that generates the predictions  $\widehat{\phi}$  (68921 nodes) in order to avoid the inverse crime (Colton et Kress, 1992; Kaipio et Somersalo, 2005). Then, a 10 dB multiplicative white gaussian noise is applied to  $\widehat{\phi}$  at the nodes of the sensors to simulate the noise inherent to experimental devices, i.e  $\widehat{\phi}_{\text{noisy}} = \widehat{\phi} (1 + 0.1 \times m)$  where  $m \sim \mathcal{N}(0, 1)$ . Synthetic data, state and control meshes can be seen in Figure 5.8. It can be seen from Figure 5.8 that 6 sources and 24 sensors,  $0.5 \text{ cm}^2$  each

one (containing 16 discretization points), are used for this particular test. A reduction of the control-space dimension is employed to improve the quality of the reconstructions obtained with the BFGS algorithm (Dubot *et al.*, 2015b) : the mesh associated to the radiative properties  $\kappa$  and  $\sigma$  contains only 9261 nodes. Note that the total number of complex measurements is equal to 2304.

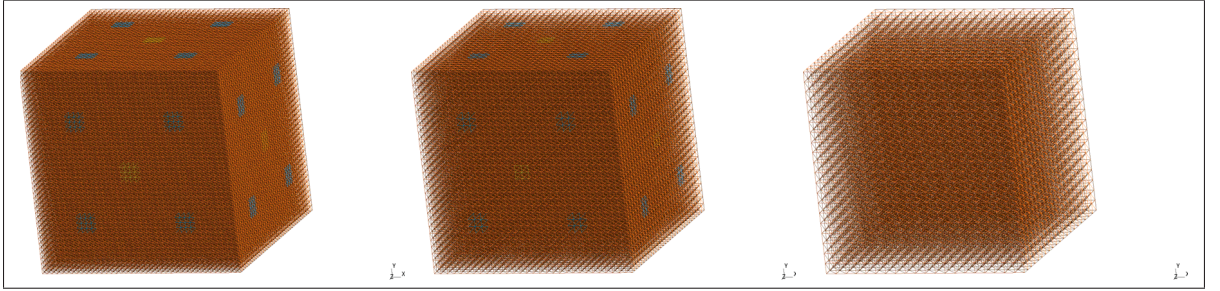


FIG. 5.8 Synthetic data mesh : 132651 nodes (left) ; mesh of the state and adjoint variable : 68921 nodes (center) ; mesh associated to radiative properties  $\kappa$  and  $\sigma$  : 9261 nodes (right).

Figure 5.9 presents the evolutions, with respect to iterations, of deviation and correlation factors, respectively to both parameters. It is seen from this Figure that, when compared to ordinary  $L^2$  gradients, the use of space-dependent Sobolev gradients decreases deviation factors and increases correlation factors. The optimum regularization parameter is about 10 in this particular case. Figure 5.10 corroborates these conclusions : the best reconstructions, in terms of smoothness and accuracy, are obtained for a value of the regularization parameter,  $\nu$ , equal to 10. Finally, remark that Figure 5.9 allows to observe the establishment of a cross-talk effect (Arridge et Lionheart, 1998; McBride *et al.*, 2000) : for  $\nu = +\infty$  and  $\nu = 70$ , it is seen that the deviation factor  $e_1^\kappa$  decreases with the iterations, while the deviation factor  $e_1^\sigma$  increases. In these two cases, values of the cost function decrease with the iterations, while an improvement of the absorption coefficient reconstruction and a deterioration of the reduced scattering coefficient reconstruction are observed. This effect is prevented by the use of space-dependent Sobolev gradients with an appropriate regularisation parameter weight.

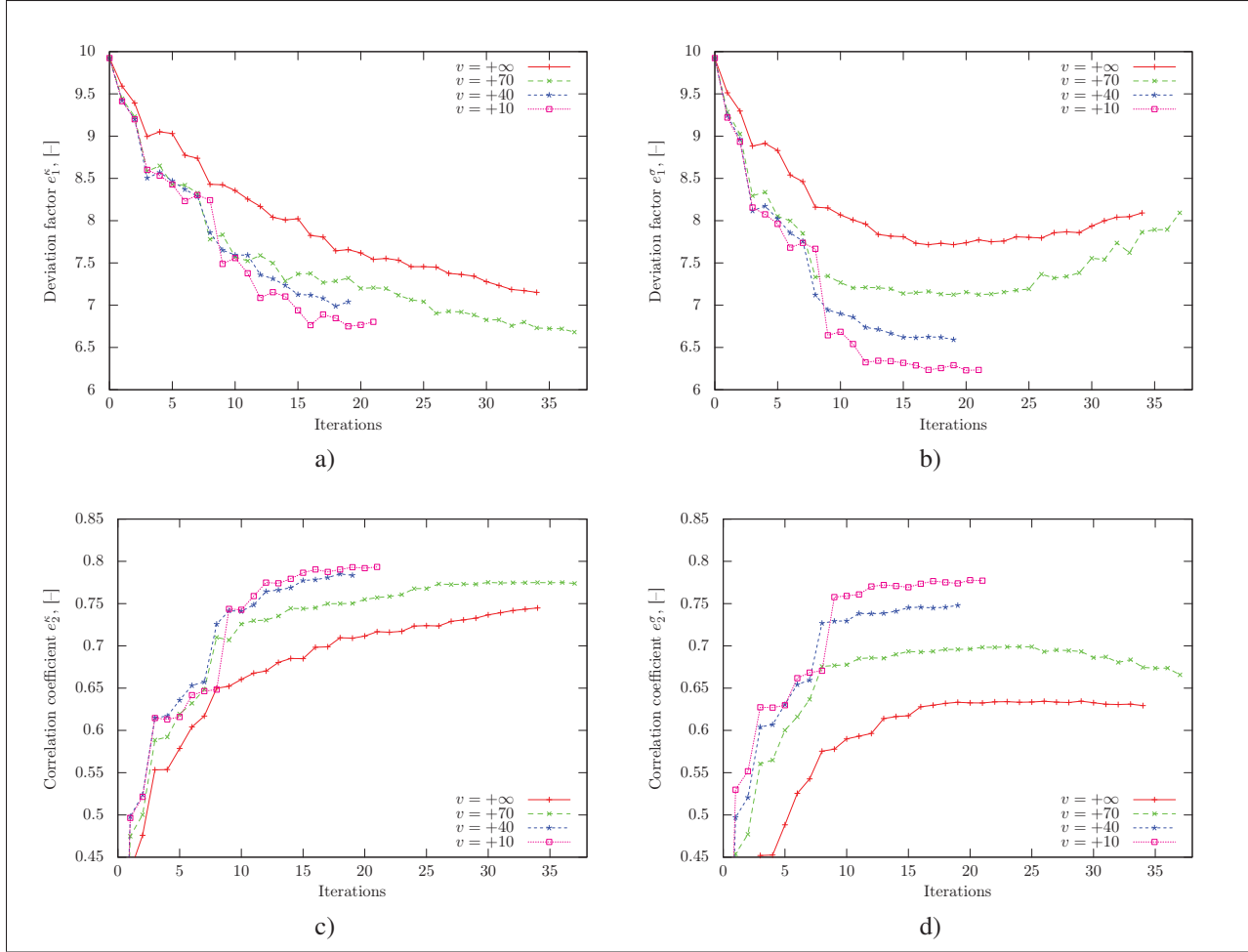


FIG. 5.9 10 dB. Evolution of errors  $e_1$  (top) and  $e_2$  (bottom) for  $\kappa$  (left) and  $\sigma$  (right), with respect to iterations, for  $\nu = +\infty, 70, 40$  and 10. The algorithms stops at iteration  $k$  when  $\frac{|j^k - j^{k-1}|}{j^{k-1}} < 10^{-3}$ .

## 5.5 Conclusion

In this paper, inverse models based on the BFGS algorithm have been developed to solve optical tomography problems based on two different forward models : the bi-dimensional steady-state radiative transfer equation and the three-dimensional frequency diffuse approximation. A Sobolev filter function was defined and space-dependent Sobolev inner products were used when extracting the cost function gradients instead of the  $L^2$ -inner product. Two test cases consisting in reconstructing discontinuous radiative properties in a square and a cube have been consid-

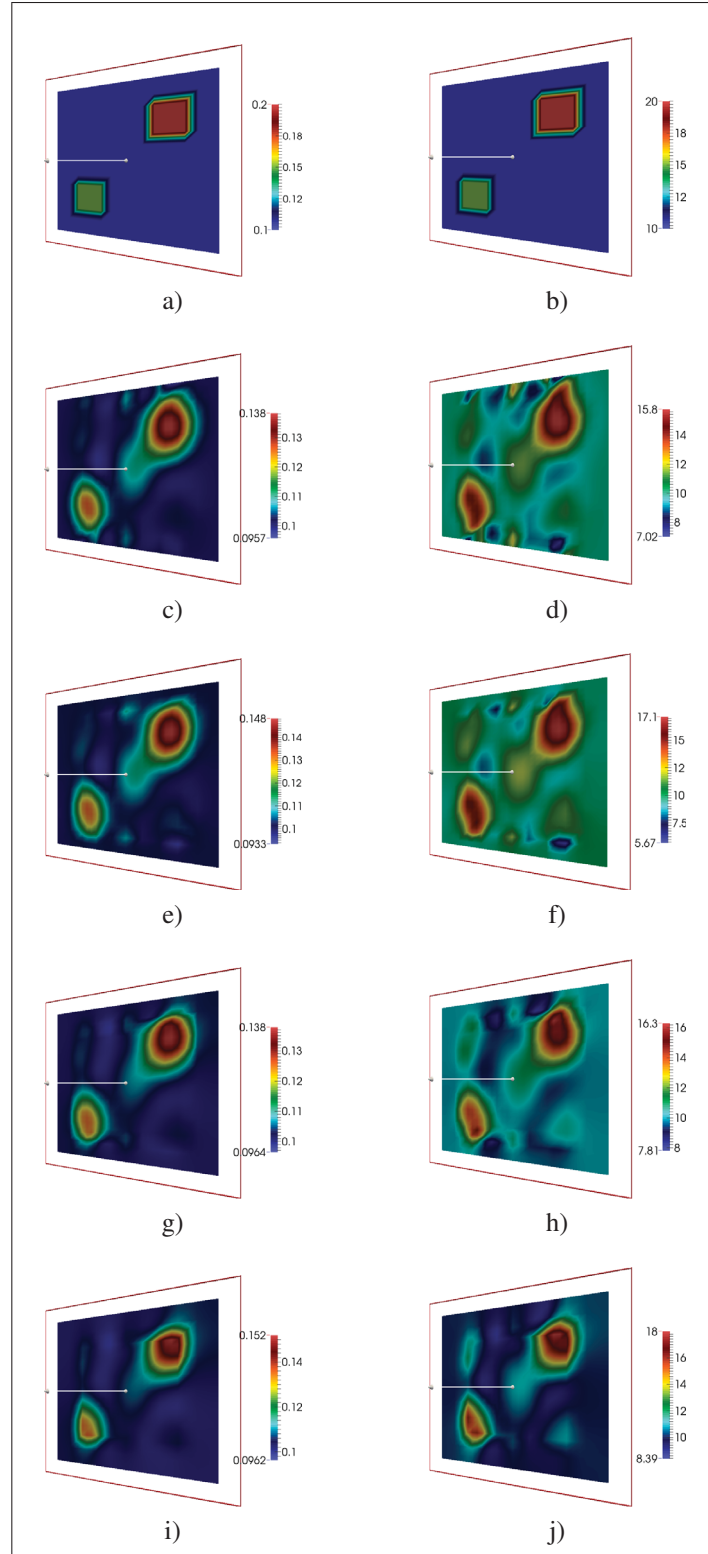


FIG. 5.10 10 dB. Reconstructions obtained in the plane passing through the center of the cube of which the normal vector  $\mathbf{n}$  equals  $(1, -1, 0)^\top$ , for  $\kappa$  (left) and  $\sigma$  (right). 1st row : targets, then :  $v = +\infty$ , 70, 40 and 10, respectively.

red. Numerical results obtained have shown that, for all levels of noise and all values of the regularization parameter tested, the use of Sobolev gradients drastically enhances the quality of the reconstructions. Also, it has been observed that the space-dependent Sobolev gradients can allow to avoid the establishment of a crosstalk effect in the case of the frequency three-dimensional approximation diffuse based inversion.

Generally speaking, due to the simplicity of its implementation and its effectiveness in filtering noise measurements locally, which provides better reconstructions, space-dependent Sobolev gradients appear to be particularly attractive in the context of inverse problems. In a near future, the anisotropy factor will be integrated in the inverse problem based on the radiative transfer equation and the use of a logarithmic cost function will be considered as suggest our more recent results. Later, inversion numerical algorithms of the three-dimensional radiative transfer equation and diffuse approximation model will be integrated in a global numerical code. The latter will be based on the BFGS algorithm with the use of adimensionalisation of the radiative properties, reduction of the control space dimension, space-dependent Sobolev gradients and a wavelet multi-scale approach, as developed in an other work.

## CHAPITRE 6

### RÉSULTATS COMPLÉMENTAIRES

#### 6.1 Méthode inverse multi-échelle basée sur les ondelettes de Haar pour le modèle direct des sources collimatées

Dans cette section, nous appliquons la méthode inverse multi-échelle développée au chapitre 4 pour des sources diffuses, au cas du modèle direct des sources collimatées. Le problème inverse à résoudre est donc celui présenté au chapitre 2. L'optimiseur utilisé pour minimiser la fonction de coût, éq. (2.13), est le BFGS tel que présenté dans les chapitres 2 et 4. Les reconstructions obtenues à l'aide de l'algorithme de BFGS ordinaire sont comparées à celles obtenues avec la méthode inverse multi-échelle développée ci-après.

La Figure 6.1 représente la configuration du test : le domaine  $\Omega$  est le carré  $[-4, 4]^2$ , 12 sources et 16 détecteurs équirépartis le long de la frontière  $\partial\Omega$  sont utilisés. Les paramètres physiques du modèle direct, éq. (2.1), sont donnés par :  $\omega = 100 \times 10^6 \text{ s}^{-1}$ ,  $c = (3 \times 10^{10}/1, 4) \text{ cm.s}^{-1}$  et  $q_l = 0, 1 \text{ W.cm}^{-2}$  pour tout  $l = 1, \dots, 12$ . La discrétisation régulière du maillage de la variable d'état  $\varphi$  est constituée de  $128 \times 128$  nœuds ( $N = -7$ ) tandis que celle des données synthétiques est constituée de  $255 \times 255$  nœuds. Enfin, les propriétés cibles sont représentées sur la Figure 6.2 (1<sup>re</sup> ligne).

L'application de la méthode multi-échelle à ce nouveau modèle direct est plus simple que pour le cas présenté au chapitre 4. En effet, une seule mesure étant collectée par détecteur, les projections unidimensionnelles sur bases d'ondelettes des prédictions et données synthétiques aux différentes échelles effectuées dans les algorithmes 4.2 et 4.3 ne sont pas applicables. Ainsi, les seules modifications de l'algorithme de BFGS, Alg. 4.1, à effectuer pour obtenir le développement de la méthode multi-échelle concernent les projections bidimensionnelles sur bases d'ondelettes des fonctions  $\varphi$ ,  $\varphi^*$ ,  $\partial_{r_1}(\varphi)$ ,  $\partial_{r_2}(\varphi)$ ,  $\partial_{r_1}(\varphi^*)$  et  $\partial_{r_2}(\varphi^*)$  lors du calcul du gradient de la fonction de coût à une échelle donnée. L'échelle maximale, déterminée par le

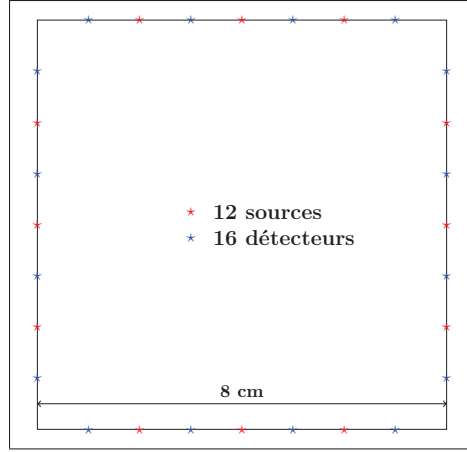


FIG. 6.1 Configuration des tests des sections 6.1 et 6.2.

nombre de points de mesures par détecteur dans le cas du chapitre 4, est à présent laissée au choix de l'utilisateur.

Comme dans les résultats numériques présentés au chapitre 4, l'échelle originale du problème inverse est  $N = -7$ . Dans la présente étude, nous avons choisi l'échelle maximale  $M = -3$  et l'ondelette de Haar est utilisée. La méthode inverse multi-échelle consiste alors en l'application de l'algorithme de BFGS de l'échelle -3 à l'échelle -7 où le gradient de la fonction de coût est calculé à l'échelle correspondante tel que précisé ci-dessus. Le passage d'une échelle donnée à l'échelle inférieure s'effectue à l'aide des critères d'arrêts suivants :  $N^{-3} = 20$ ,  $N^{-4} = 40$ ,  $N^{-5} = 60$ ,  $N^{-6} = 80$  et  $N^{-7} = 100$ , où  $N^j$  est le nombre maximal d'itérations autorisé pour l'inversion à l'échelle  $j$ ,  $j = -7, \dots, -3$ . De plus, un critère d'arrêt sur la stabilisation de la fonction de coût est ajouté : l'algorithme passe à l'échelle inférieure lorsque  $\frac{|j(\gamma^k) - j(\gamma^{k-1})|}{j(\gamma^k)} \leq 10^{-2}$ , où  $\gamma^k$  représente les variables de contrôle à l'itération  $k$ . Rappelons également que la solution du problème d'optimisation à une échelle donnée est utilisée comme initialisation du problème d'optimisation à l'échelle inférieure. Enfin, concernant la méthode de BFGS ordinaire, le même critère d'arrêt sur la stabilisation de la fonction de coût est utilisé et le nombre maximal d'itérations autorisé est fixé à 300 afin d'avoir, dans le pire des cas, un même nombre d'itérations pour les deux méthodes.



Tout d'abord, une étude comparable à celle menée dans le chapitre 2 sur les effets de la régularisation par reparamétrisation des variables de contrôle est réalisée. Les Figures 6.2 et 6.3 présentent les reconstructions obtenues par l'algorithme de BFGS ordinaire pour des données synthétiques bruitées à 20 dB et 10 dB, respectivement, et pour une paramétrisation fine (égale à celle de la variable d'état) et grossière des variables de contrôle (2<sup>e</sup> et 3<sup>e</sup> lignes, respectivement). Comme en attestent les erreurs de reconstruction  $e_1$  et  $e_2$  données sur ces figures, une nette amélioration des résultats est obtenue par l'utilisation de cette régularisation, en accord avec les conclusions du chapitre 2. D'ailleurs, cette régularisation permet la convergence de la méthode inverse dans le cas d'un bruit de 10 dB (Fig. 6.3, 2<sup>e</sup> ligne, divergence des reconstructions en absence de régularisation).

Les Figures 6.4 et 6.5 comparent les reconstructions obtenues avec l'algorithme de BFGS ordinaire et utilisation de la méthode de régularisation par paramétrisation (2<sup>e</sup> ligne) à celles obtenues par la méthode inverse multi-échelle (3<sup>e</sup> ligne) en présence d'un bruit de 20 dB et 10 dB, respectivement. Remarquons que nous avons ajouté l'étape de régularisation par paramétrisation à la méthode inverse multi-échelle en projetant les gradients obtenus aux différentes échelles sur la paramétrisation grossière. Les erreurs de reconstruction indiquées dans ces figures montrent que la méthode multi-échelle permet d'atteindre un niveau de précision des reconstructions très supérieur à celui obtenu par l'algorithme de BFGS ordinaire.

Enfin, les Figures 6.6 et 6.7 comparent les reconstructions obtenues par les deux méthodes proposées lorsque les propriétés radiatives initiales sont sur-estimées de 5 % par rapport aux valeurs des propriétés radiatives en arrière-plan du milieu  $\Omega$  et pour un bruit de 20 dB et 10 dB, respectivement. Nous observons comme précédemment une très grande amélioration de la qualité des résultats par l'utilisation de la méthode multi-échelle.

Pour finir, nous pouvons remarquer que l'amélioration des reconstructions obtenues par l'utilisation de la méthode multi-échelle est encore plus probante dans le cas présent du modèle direct pour des sources collimatées que pour celui considérant des sources diffuses tel qu'utilisé dans le chapitre 4.

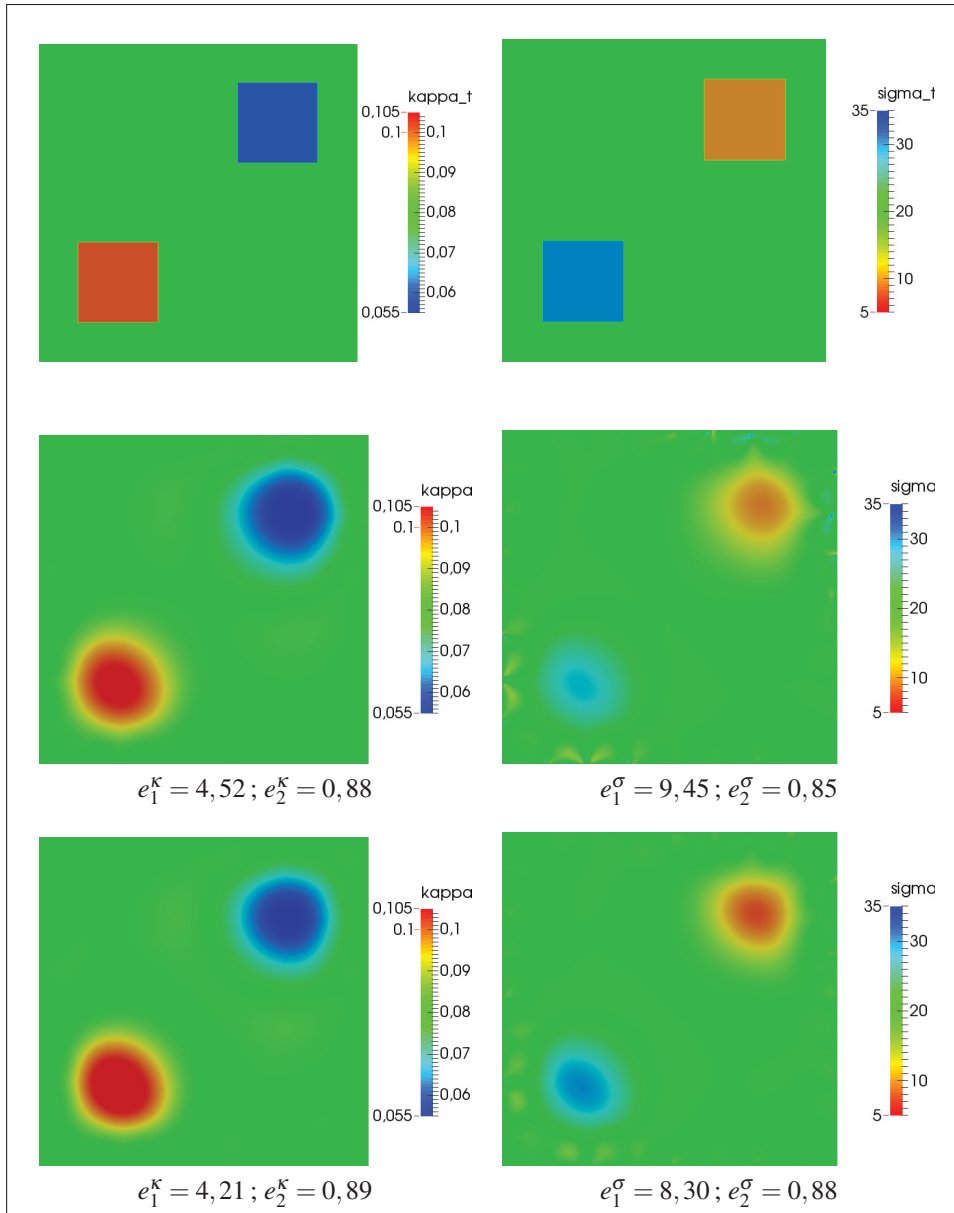


FIG. 6.2 Reconstructions  $\kappa$  et  $\sigma$ . SNR = 20 dB.  $\kappa_{ap} = 0,08$ ,  $\sigma_{ap} = 20$ . 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS, maillage composé de  $128 \times 128$  nœuds. 3<sup>e</sup> ligne : méthode de BFGS, maillage composé de  $80 \times 80$  nœuds. Gauche :  $\kappa$  ; Droite :  $\sigma$ .

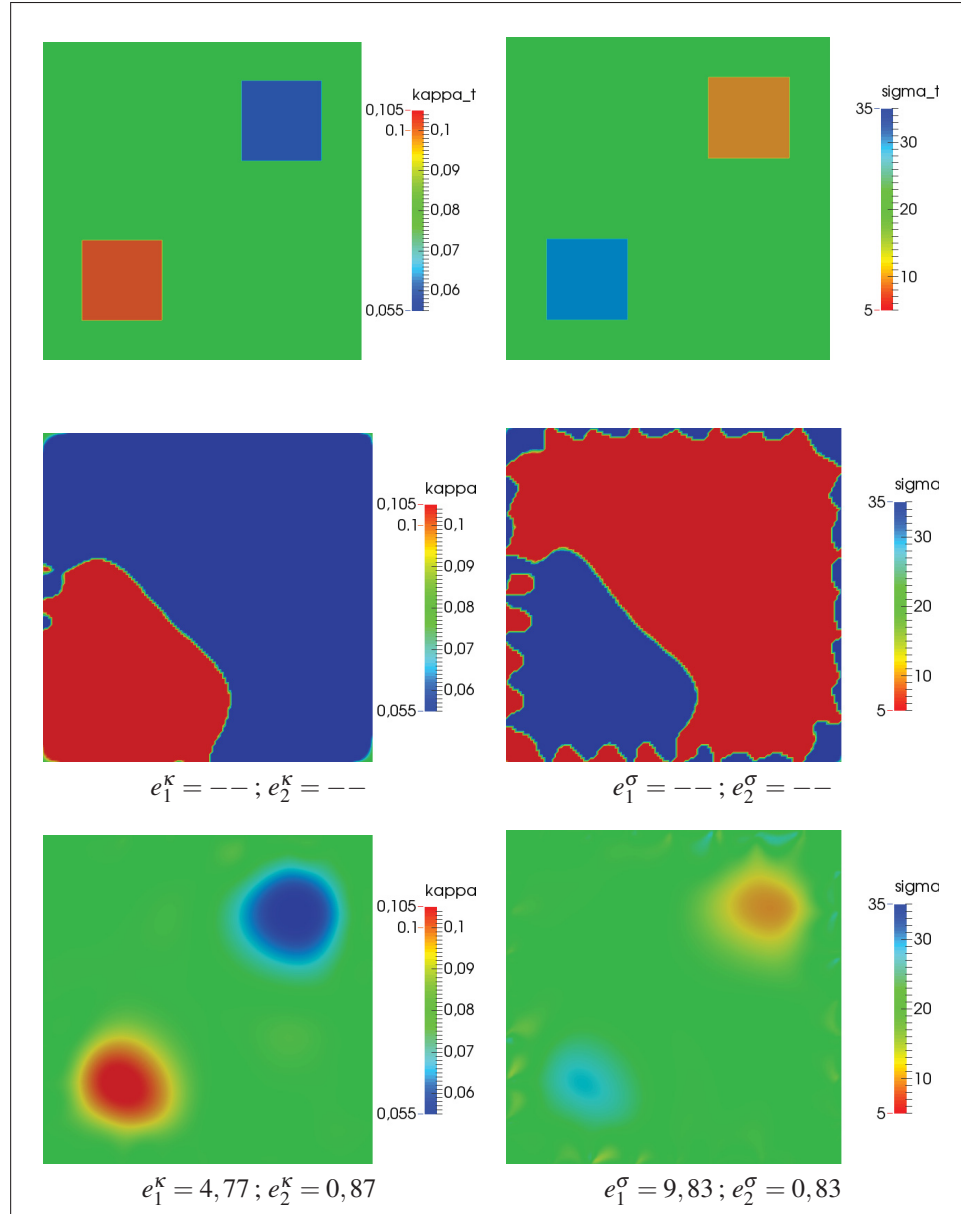


FIG. 6.3 Reconstructions  $\kappa$  et  $\sigma$ . SNR = 10 dB.  $\kappa_{ap} = 0,08$ ,  $\sigma_{ap} = 20$ . 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS, maillage composé de  $128 \times 128$  nœuds. 3<sup>e</sup> ligne : méthode de BFGS, maillage composé de  $80 \times 80$  nœuds. Gauche :  $\kappa$  ; Droite :  $\sigma$ .

## 6.2 Crosstalk entre les paramètres

Nous abordons dans cette section un des verrous principal de la tomographie optique, à savoir le phénomène dit de *crosstalk* entre les paramètres. Plusieurs auteurs ont déjà fait état de ce phé-

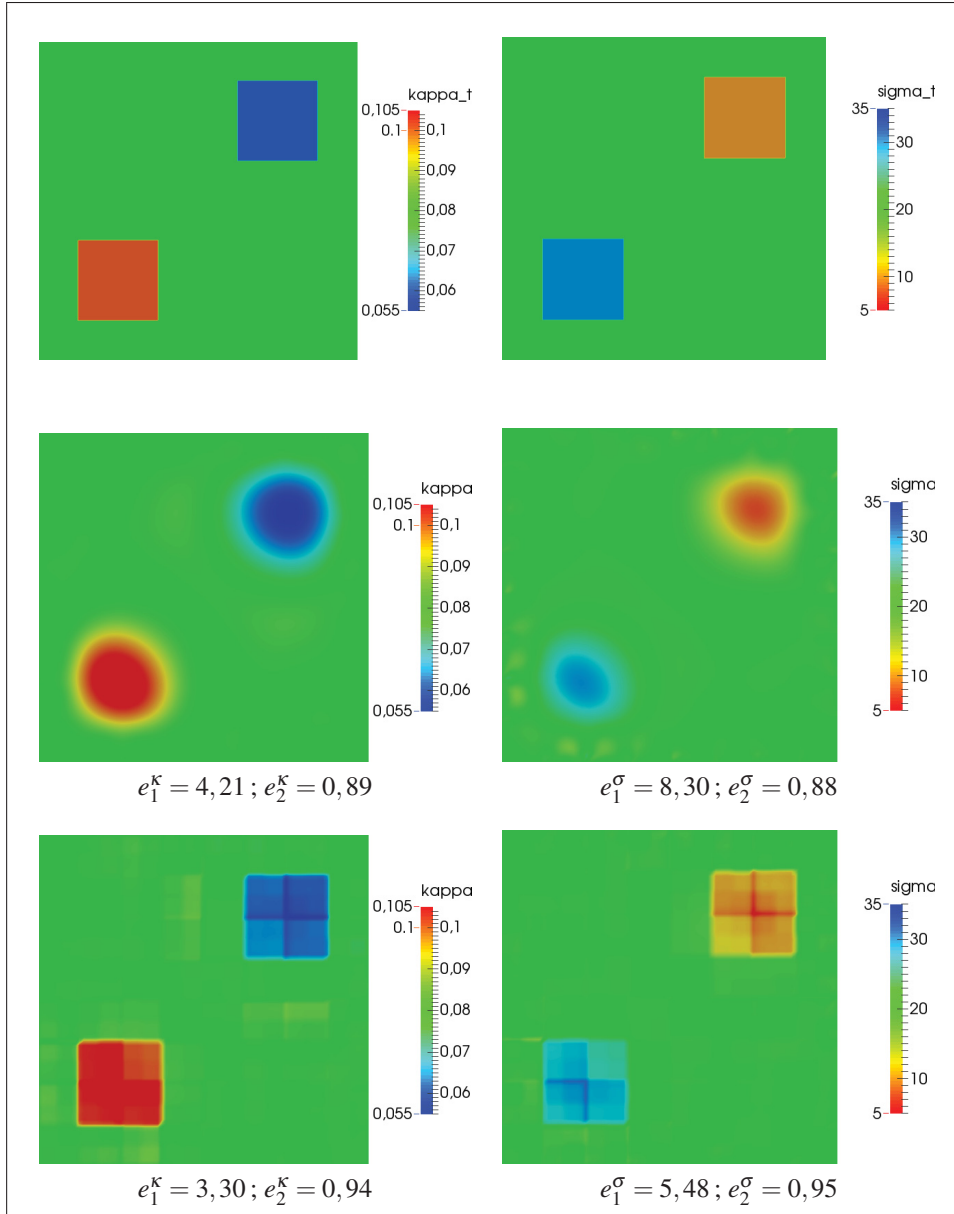


FIG. 6.4 Reconstructions  $\kappa$  et  $\sigma$ . SNR = 20 dB.  $\kappa_{ap} = 0,08$ ,  $\sigma_{ap} = 20$ . Maillage composé de  $80 \times 80$  nœuds. 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS. 3<sup>e</sup> ligne : méthode multi-échelle. Gauche :  $\kappa$  ; Droite :  $\sigma$ .

nomène entre les coefficients d'absorption et de diffusion réduit dans le cas de la TO basée sur l'AD (Arridge et Lionheart, 1998; Schweiger et Arridge, 1999; McBride *et al.*, 2000, 2001b; Pei *et al.*, 2001; Xu *et al.*, 2002). Ce phénomène se rapporte au cas où, par exemple, des varia-

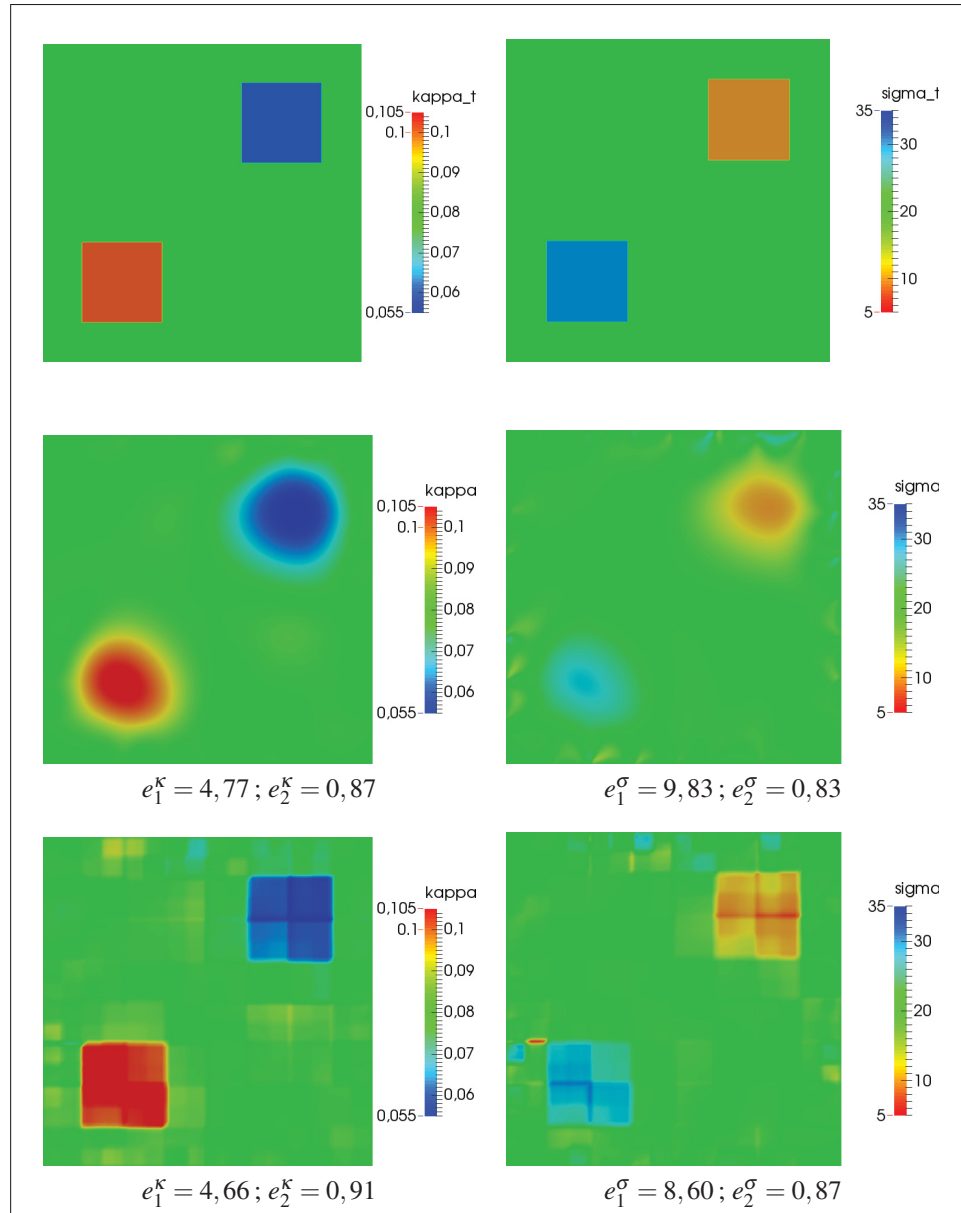


FIG. 6.5 Reconstructions  $\kappa$  et  $\sigma$ . SNR = 10 dB.  $\kappa_{ap} = 0,08$ ,  $\sigma_{ap} = 20$ . Maillage composé de  $80 \times 80$  nœuds. 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS. 3<sup>e</sup> ligne : méthode multi-échelle. Gauche :  $\kappa$  ; Droite :  $\sigma$ .

tions spatiales du coefficient d'absorption apparaissent également dans l'image reconstruite du coefficient de diffusion réduit, ou *vice versa*.

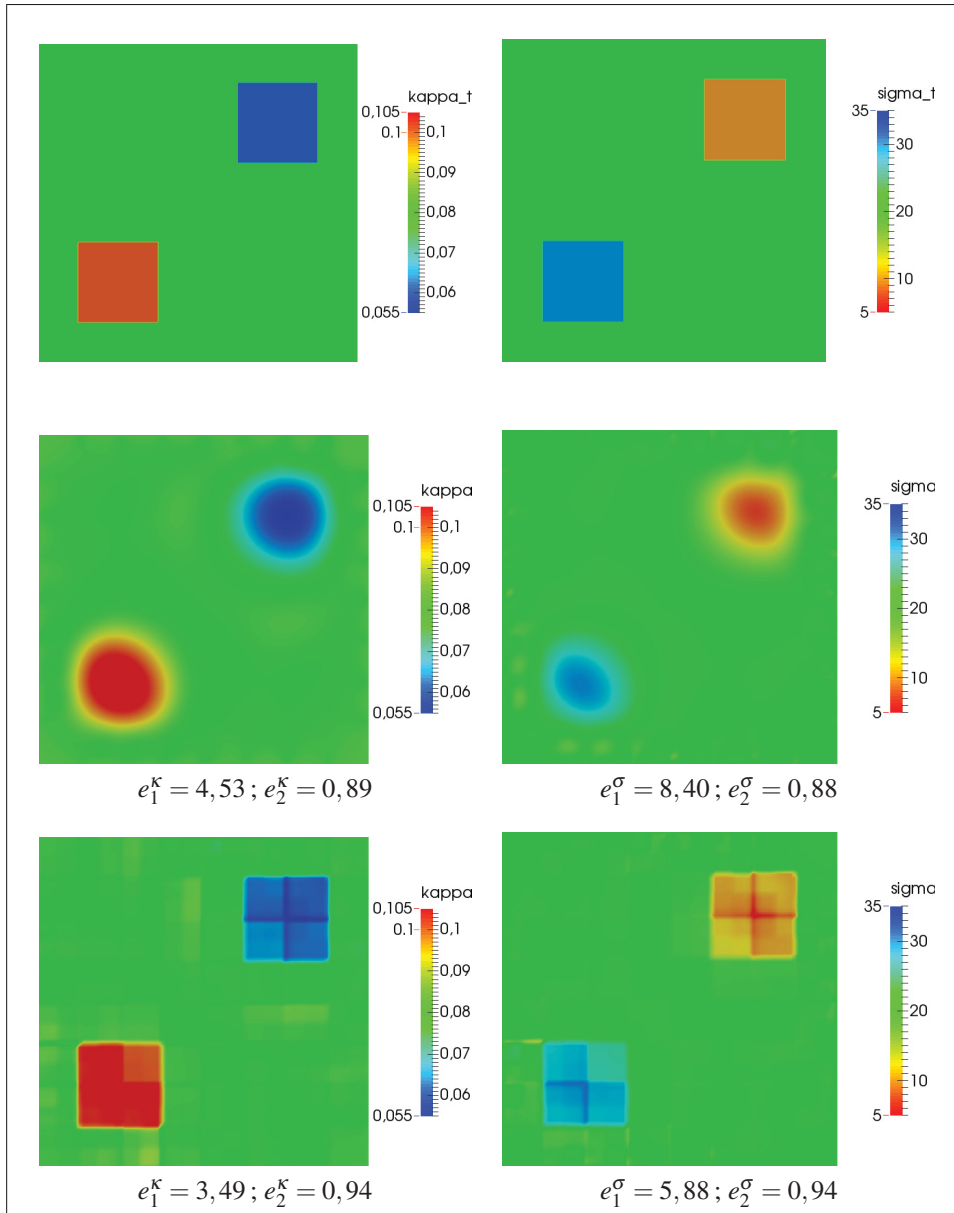


FIG. 6.6 Reconstructions  $\kappa$  et  $\sigma$ . SNR = 20 dB.  $\kappa_{ap} = 0,084$ ,  $\sigma_{ap} = 21$ . Maillage composé de  $80 \times 80$  nœuds. 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS. 3<sup>e</sup> ligne : méthode multi-échelle. Gauche :  $\kappa$  ; Droite :  $\sigma$ .

Ce phénomène peut être compris en examinant les expressions des gradients de la fonction de coût par rapport à chacun des paramètres, éq. (2.41). En effet, nous observons la corrélation

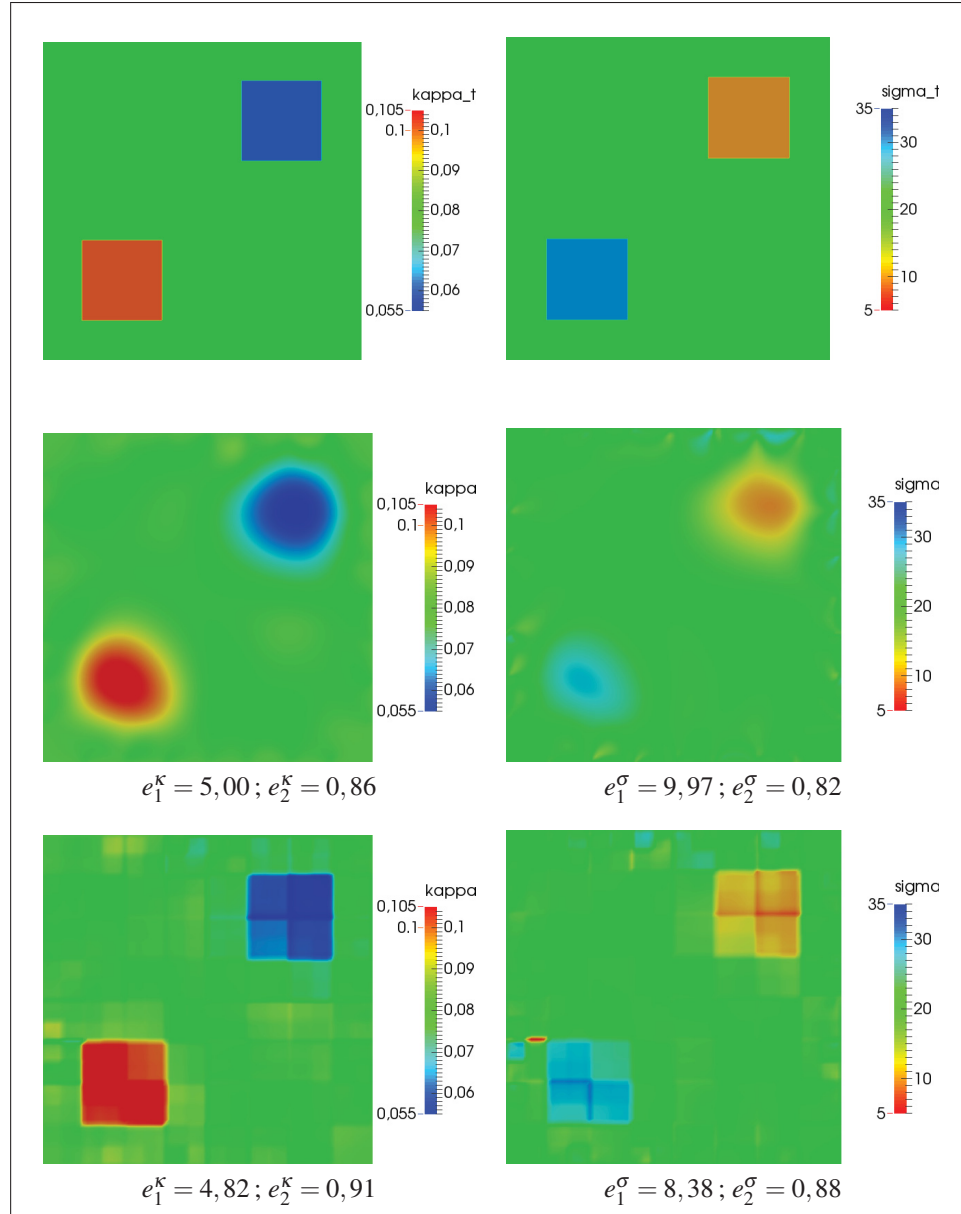


FIG. 6.7 Reconstructions  $\kappa$  et  $\sigma$ . SNR = 10 dB.  $\kappa_{ap} = 0,084$ ,  $\sigma_{ap} = 21$ . Maillage composé de  $80 \times 80$  nœuds. 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS. 3<sup>e</sup> ligne : méthode multi-échelle. Gauche :  $\kappa$  ; Droite :  $\sigma$ .

suivante entre les deux gradients :

$$\begin{aligned} \nabla_{\kappa} j &= -n_{\Omega} D^2 \operatorname{Re}(\nabla \varphi \cdot \nabla \varphi^*) + \operatorname{Re}(\varphi \varphi^*) \\ \Rightarrow \nabla_{\kappa} j &= \nabla_{\sigma} j + \operatorname{Re}(\varphi \varphi^*) \end{aligned} \quad (6.1)$$

Ainsi, imaginons par exemple une hétérogénéité du coefficient de diffusion réduit en un endroit tandis que le coefficient d'absorption est uniforme en ce même endroit. Alors les valeurs du gradient de la fonction de coût par rapport à  $\sigma$  seront probablement grandes en cet endroit, en valeur absolue. Comme le gradient de la fonction de coût par rapport à  $\sigma$  apparaît dans celui par rapport à  $\kappa$ , on en déduit que ce dernier pourra possiblement être corrompu, entraînant une hétérogénéité dans la reconstruction du coefficient d'absorption n'ayant pas lieu d'être. Finalement, nous pouvons remarquer que l'importance du *crosstalk* dans la reconstruction du coefficient d'absorption dû à la diffusion réduite, ou *vice versa*, résulte d'une compétition entre les expressions  $\text{Re}(\varphi\varphi^*)$  et  $n_\Omega D^2 \text{Re}(\nabla\varphi \cdot \nabla\varphi^*)$ .

Nous comparons dans la suite les comportements des méthodes de BFGS ordinaire et multi-échelle face au phénomène de *crosstalk*. La configuration du test est la même que celle utilisée dans la section précédente (section 6.1). Dans le premier test, nous examinons le *crosstalk* dans la reconstruction du coefficient d'absorption dû au coefficient de diffusion réduit. La 1<sup>re</sup> ligne de la Figure 6.8 présente les propriétés cibles : on observe une hétérogénéité du coefficient de diffusion réduit à l'endroit où les valeurs du coefficient d'absorption restent uniformes (coin sud-ouest). La 2<sup>e</sup> ligne présente les reconstructions obtenues avec l'algorithme de BFGS ordinaire tandis que la 3<sup>e</sup> ligne présente les reconstructions obtenues avec la méthode inverse multi-échelle. Les erreurs de reconstruction ainsi que l'observation visuelle indiquent des reconstructions plus précises pour la méthode multi-échelle. Cependant, nous constatons un très fort *crosstalk* entre les paramètres pour les deux méthodes : l'inclusion sud-ouest du coefficient de diffusion réduit apparaît clairement dans la reconstruction du coefficient d'absorption.

Dans le second test, nous examinons le *crosstalk* inverse. La 1<sup>re</sup> ligne de la Figure 6.9 présente les propriétés cibles : une hétérogénéité du coefficient d'absorption est présente là où les valeurs du coefficient de diffusion réduit restent uniformes (coin sud-ouest). Comme précédemment, nous constatons que les reconstructions obtenues avec la méthode inverse multi-échelle sont beaucoup plus précises que celles obtenues avec l'algorithme de BFGS ordinaire. Nous observons cependant que le *crosstalk* dans la reconstruction du coefficient de diffusion réduit dû au coefficient d'absorption est inexistant. Notons que ce fait a également été observé lors



de l'étude du *crosstalk* entre les coefficients d'absorption et de diffusion dans le cas de la TO basée sur l'ÉTR dans (Klose et Hielscher, 2003). En revanche, nous constatons que l'absence d'inclusion dans le quart sud-ouest du coefficient de diffusion réduit rend plus difficile la reconstruction de cette inclusion absorbante, surtout pour la méthode de BFGS ordinaire. Enfin, nous pouvons déduire de ces deux études que le terme  $n_{\Omega} D^2 \text{Re}(\nabla \varphi \cdot \nabla \varphi^*)$  est dominant par rapport au terme  $\text{Re}(\varphi \varphi^*)$ .

En conclusion, bien que la méthode multi-échelle améliore les reconstructions obtenues par rapport à l'algorithme de BFGS ordinaire, celle-ci ne permet pas de contrer le phénomène de *crosstalk* entre les paramètres. À notre connaissance, seuls les travaux de (Pei *et al.*, 2001) ont permis de faire face de manière satisfaisante au phénomène de *crosstalk*. Ces travaux ont d'ailleurs fait l'objet d'un brevet (Barbour *et al.*, 2003).

### 6.3 Reconstructions tridimensionnelles

Nous présentons dans cette section les résultats d'un cas test de reconstructions simultanées des coefficients d'absorption et de diffusion réduit d'un milieu semi-transparent tridimensionnel. Le domaine considéré est un cube de côté 5 cm,  $\Omega = [0, 5]^3$ . Six sources diffuses, chacune située au centre d'une des faces du cube et formant un carré de côté 5 mm sont utilisées. Vingt-quatre détecteurs de taille identique aux sources, quatre sur chacune des faces du cube, sont aussi utilisés pour collecter les données synthétiques. La disposition des sources et détecteurs est représentée sur la Figure 6.10-a. La Figure 6.10 présente également les maillages utilisés pour la discrétisation des variables d'état ( $\varphi$ ) et adjointe ( $\varphi^*$ ) (Figure 6.10-b) et pour la discrétisation des variables de contrôle  $\kappa$  et  $\sigma$  (Figure 6.10-c). La dimension du maillage de la variable d'état est de 47 493 nœuds et 245 612 tétraèdres. Suivant les conclusions obtenues dans le cas bidimensionnel sur la régularisation par reparamétrisation des variables de contrôle, le maillage utilisé pour la discrétisation des coefficients d'absorption et de diffusion est composé de 22 945 nœuds et 112 618 tétraèdres. Notons que le logiciel Gmsh (Geuzaine et Remacle, 2009) a été utilisé pour générer ces maillages.

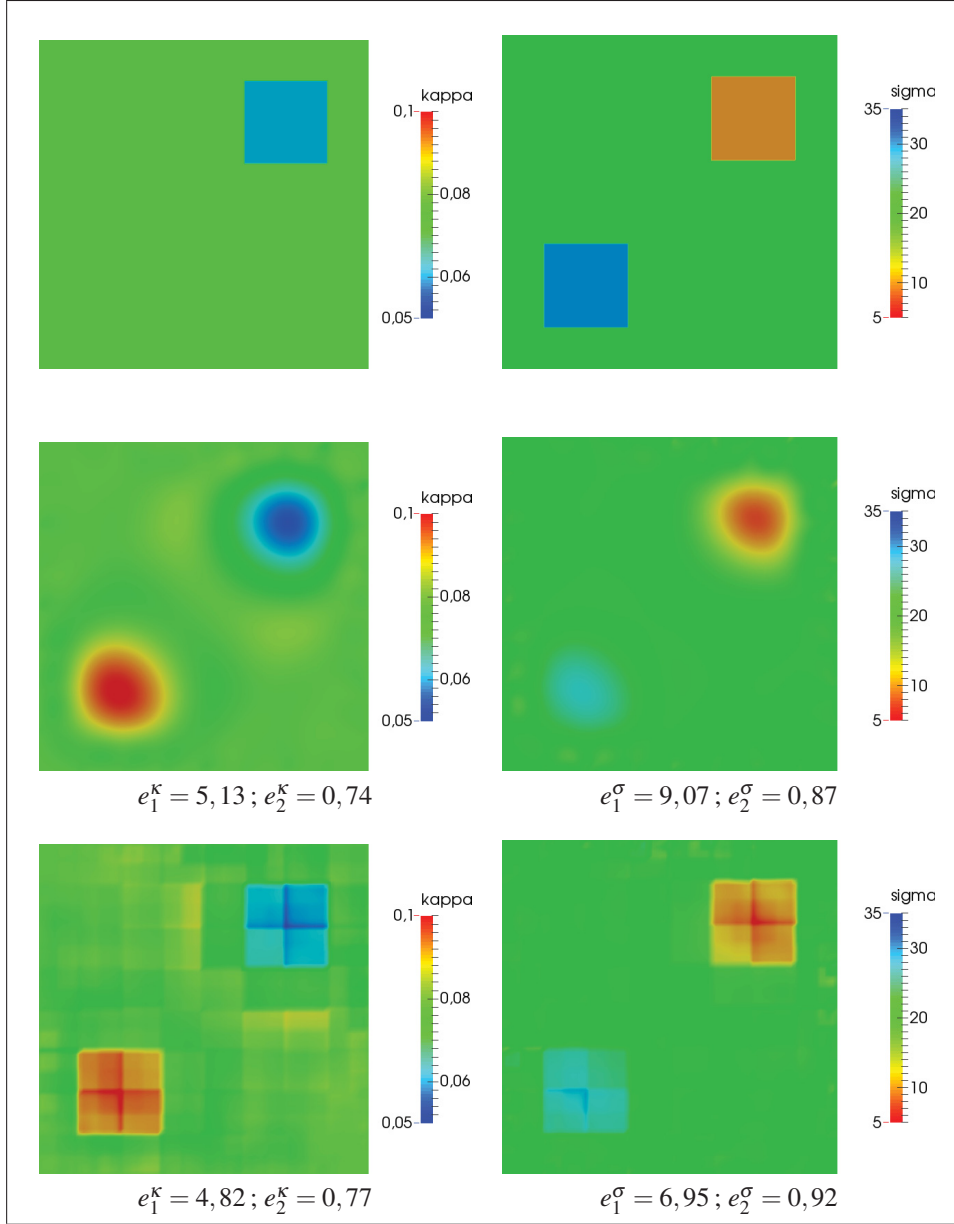


FIG. 6.8 Illustration du phénomène de *crosstalk* entre les paramètres (1).  $\text{SNR} = 20 \text{ dB}$ .  $\kappa_{ap} = 0,08$ ,  $\sigma_{ap} = 20$ . Maillage composé de  $80 \times 80$  nœuds. 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS. 3<sup>e</sup> ligne : méthode multi-échelle. Gauche :  $\kappa$ ; Droite :  $\sigma$ .

Des sources diffuses étant utilisées, le modèle direct est constitué, pour chacune des sources, des équations (1.86) et (1.87). Comme dans les chapitres 2, 3, 4 et 5, toutes les variables sont approchées sur la base des éléments finis de Lagrange  $\mathcal{P}_1$ . De même, une attention particulière

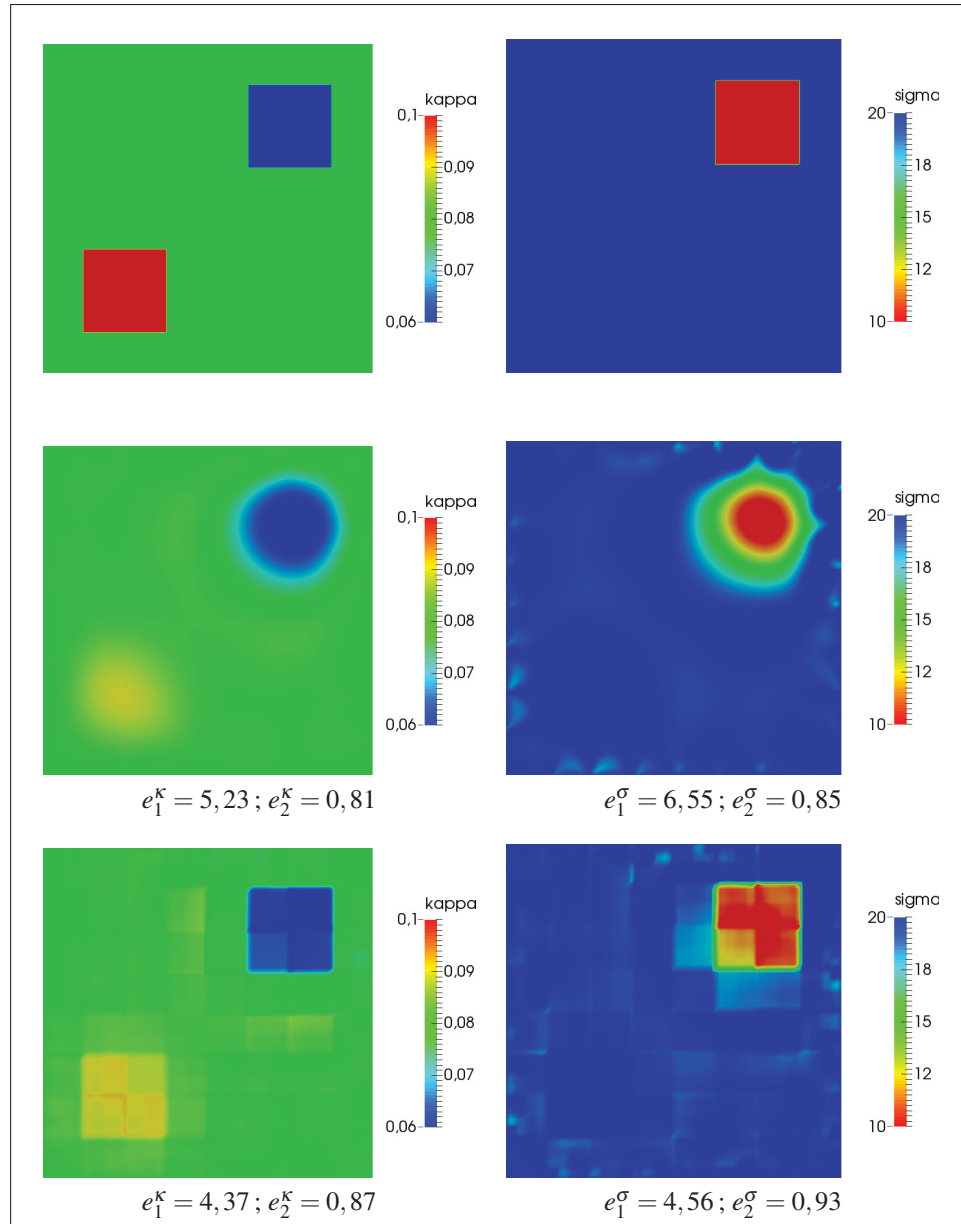


FIG. 6.9 Illustration du phénomène de *crosstalk* entre les paramètres (2).  $\text{SNR} = 20 \text{ dB}$ .  $\kappa_{ap} = 0,08$ ,  $\sigma_{ap} = 20$ . Maillage composé de  $80 \times 80$  nœuds. 1<sup>re</sup> ligne : propriétés cibles. 2<sup>e</sup> ligne : méthode de BFGS. 3<sup>e</sup> ligne : méthode multi-échelle. Gauche :  $\kappa$ ; Droite :  $\sigma$ .

a été portée à la création des données synthétiques : celles-ci sont générées sur un maillage composé de 67 103 nœuds et 340 881 tétraèdres afin d'éviter le crime inverse. Un bruit blanc

gaussien de 20 dB est ensuite ajouté à ces dernières afin de simuler le bruit inhérent aux mesures expérimentales.

L'algorithme de BFGS ordinaire tel que présenté dans les chapitres 2 et 4 est utilisé pour la minimisation de la fonction de coût pour la reconstruction des propriétés radiatives. Cependant, le critère d'arrêt considéré ici consiste uniquement en un nombre maximal d'itérations fixé à 100. De plus, la méthode de recherche linéaire employée est la méthode de recherche "inexacte" dû à Fletcher (Fletcher, 1980, p. 126) avec les paramètres  $\rho_F = 0,1$ ,  $\sigma_F = 0,7$  tel que recommandé par (Antoniou et Lu, 2007, p. 113). Enfin, les variables de contrôle adimensionnelles  $\varkappa$  et  $\varsigma$  sont initialisées à l'unité tandis que les fonctions *a priori* sont fixées à  $\kappa_{\text{ap}} = 0,08$  et  $\sigma_{\text{ap}} = 20$ . Les propriétés cibles sont quant à elles définies par :

$$\begin{aligned} \kappa(\mathbf{r}) &= 0,06 \text{ cm}^{-1}, \quad \sigma(\mathbf{r}) = 10 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega_1, & \Omega_1 &= [3,25,4,25]^3 \text{ cm}^3 \\ \kappa(\mathbf{r}) &= 0,10 \text{ cm}^{-1}, \quad \sigma(\mathbf{r}) = 30 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega_2, & \Omega_2 &= [0,75,1,75]^3 \text{ cm}^3 \\ \kappa(\mathbf{r}) &= 0,08 \text{ cm}^{-1}, \quad \sigma(\mathbf{r}) = 20 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega \setminus (\Omega_1 \cup \Omega_2), & \Omega &= [0,5]^3 \text{ cm}^3 \end{aligned} \quad (6.2)$$

La Figure 6.11 présente une coupe diagonale des reconstructions obtenues lorsque le produit scalaire usuel  $L^2(\Omega)$  est utilisé pour extraire le gradient de la fonction de coût ( $\vartheta^S = 0$ ) et lorsque le produit scalaire de Sobolev pondéré  $H^{1,\vartheta^S}(\Omega)$  tel que défini dans le chapitre 2 à la sous-section 2.5.6, avec  $\vartheta^S = 10^{-2}$ , est utilisé pour extraire le gradient de la fonction de coût. On observe que les deux inclusions sont bien retrouvées pour les deux propriétés radiatives, autant en terme de localisation que de valeurs reconstruites, avec cependant une très nette amélioration des résultats avec l'utilisation des gradients de Sobolev. En comparant avec les résultats bidimensionnels obtenus à la sous-section 2.5.6 du chapitre 2 pour des données bruitées à 20 dB, nous pouvons remarquer que l'utilisation des gradients de Sobolev semble plus pertinente dans le cas tridimensionnel que dans le cas bidimensionnel.

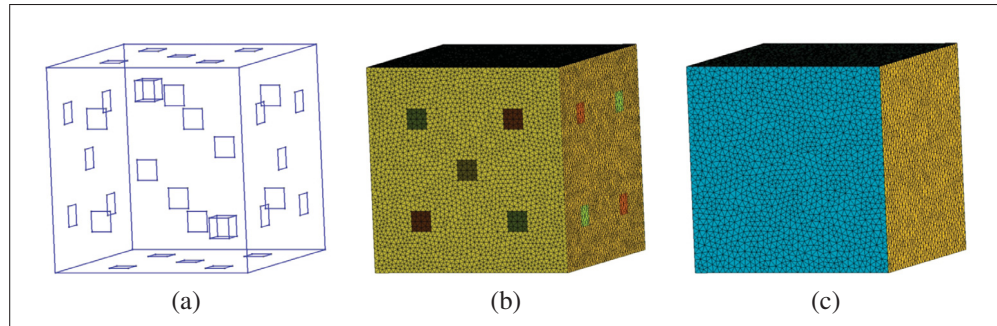


FIG. 6.10 (a) Géométrie utilisée pour le calcul des données synthétiques, contenant les deux inclusions ; (b) maillage de la variable d'état ; (c) maillage des variables de contrôle.

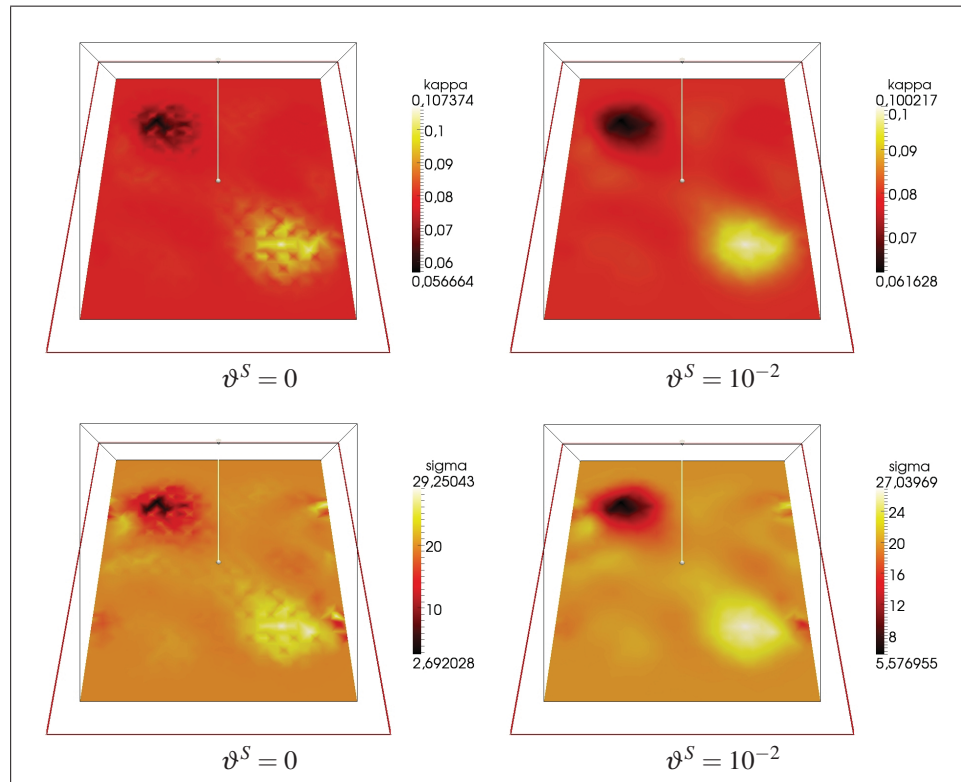


FIG. 6.11 Coupes diagonales des reconstructions des propriétés radiatives tridimensionnelles pour deux valeurs du paramètre de régularisation de Sobolev  $\vartheta^S$ . Haut : coefficient d'absorption ; Bas : coefficient de diffusion réduit.



## CHAPITRE 7

### DISCUSSIONS DES RÉSULTATS

L'objectif de cette thèse consistait à développer des outils d'inversion robustes pour la résolution de problèmes de TOD dans le domaine fréquentiel. Bien qu'il ne nous a pas été possible de confronter les algorithmes développés à l'expérimentation, nous avons pris soin tout au long de ce travail d'utiliser des données synthétiques simulant correctement des mesures expérimentales. Pour ce faire, les données ont été générées sur un maillage plus fin que celui utilisé pour calculer les prédictions. Cette étape a pour but d'éviter, ce qu'on appelle communément dans le domaine de l'inversion, le crime inverse. Les données sont ensuite projetées dans l'espace d'approximation de la variable d'état. Un bruit blanc gaussien multiplicatif, de rapport signal sur bruit fixé, est enfin ajouté à ces données afin de simuler le bruit inhérent aux mesures expérimentales.

Pour répondre à l'objectif, la première étape a consisté en l'implémentation d'un code numérique performant de résolution du modèle de l'approximation diffuse bidimensionnel et tridimensionnel dans le régime fréquentiel pour des sources diffuses ou collimatées. La méthode des éléments finis a été choisie notamment pour sa facilité de prise en compte des géométries complexes. Une validation numérique de cette résolution est présentée dans l'annexe III. Les performances du code développé reposent en particulier sur celles du solveur MUMPS utilisé pour l'inversion du système matriciel. D'ailleurs, il convient de noter que la puissance de la méthode des éléments finis est complètement exploitée dans la structure même de notre code d'inversion. En effet, que ce soit pour le calcul de la fonction de coût ou de son gradient, l'assemblage des matrices de raideur associées aux formes bilinéaires pour les modèles direct et adjoint, étape la plus coûteuse en temps de calcul pour le solveur choisi, n'est réalisée qu'une seule fois peu importe le nombre de sources utilisées, économisant ainsi des temps CPU considérables.

La seconde étape a consisté à développer des algorithmes d'optimisation robustes et performants pour minimiser la fonction de coût, qui intègre les prédictions et les mesures, et per-

mettent ainsi la reconstruction des distributions spatiales des coefficients d'absorption et de diffusion réduit. Cette seconde étape a été de pair avec la troisième étape qui consistait à développer des méthodes de régularisation adaptées au caractère mal-posé du problème inverse. Ces questions ont d'abord été abordées dans le cas bidimensionnel.

En premier lieu, un code d'inversion basé sur l'algorithme de BFGS a été développé. Cette méthode requiert en particulier la connaissance des gradients de la fonction de coût et une méthode de recherche linéaire afin de déterminer le pas de descente entre chacune des itérations. La méthode de l'état adjoint, optimale pour des problèmes d'optimisation de grande taille, a été utilisée pour le calcul des gradients de la fonction de coût. Les développements détaillés du calcul de ces gradients sont présentés en annexe I. En ce qui concerne les méthodes de recherche linéaire, la méthode de la section d'or, la méthode de l'interpolation quadratique et une méthode de recherche linéaire inexacte due à Fletcher ont été comparées pour l'algorithme de BFGS dans l'article de conférence présenté en annexe II. De même, le choix des fonctions de forme pour la paramétrisation des variables de contrôle a été étudié dans cet article. Les éléments finis de Lagrange  $\mathcal{P}_0$ ,  $\mathcal{P}_1$ ,  $\mathcal{P}_2$  et  $\mathcal{P}_3$  ont été utilisés et comparés. Les conclusions de cette étude sont que la paramétrisation  $\mathcal{P}_1$  et la méthode de recherche linéaire de l'interpolation quadratique fournissent les reconstructions les plus précises, pour un temps CPU fixé. En effet, parmi les différentes paramétrisations basées sur les éléments finis de Lagrange, la paramétrisation  $\mathcal{P}_1$  est celle pour laquelle le nombre de degrés de liberté correspondant est le plus faible. Choisir donc la fonction de forme  $\mathcal{P}_1$  parmi les autres fonctions de forme de Lagrange revient quelque part à régulariser implicitement le problème inverse, tout comme il a été fait par ailleurs sur le choix de finesse de maillage dans l'étape de régularisation. Pour ce qui est de la recherche linéaire, le fait que l'interpolation quadratique soit, parmi les différentes méthodes testées, celle qui permet les meilleures reconstructions en un temps CPU donné, peut s'expliquer par le fait que la fonction de coût ne s'écarte pas totalement de son caractère quadratique.

Deuxièmement, une comparaison entre les algorithmes de Gauss-Newton amorti et de BFGS a été effectuée. Deux méthodes de régularisation sont combinées pour chacun des deux algo-



rithmes, à savoir la réduction de la dimension de l'espace de contrôle basée sur le maillage et la régularisation par pénalisation de Tikhonov pour l'algorithme de Gauss-Newton amorti, et les régularisations basées sur le maillage et l'utilisation des gradients de Sobolev, uniformes ou spatialement dépendants, lors de l'extraction du gradient de la fonction coût, pour la méthode BFGS. De plus, deux stratégies visant à contrer l'effet préjudiciable de la différence d'ordre de grandeur entre les propriétés d'absorption et de diffusion réduit sur la qualité des reconstructions, à savoir une mise à l'échelle des gradients de la fonction de coût et une adimensionnalisation des variables de contrôle, sont comparées pour la méthode de BFGS. Seul l'adimensionnalisation des variables de contrôle est utilisée pour l'algorithme de Gauss-Newton amorti. Cette étude a permis de constater la supériorité de l'algorithme de BFGS par rapport à celui de Gauss-Newton amorti en ce qui concerne à la fois la qualité des reconstructions obtenues, le temps de calcul ou encore la facilité de sélection du paramètre de régularisation. La régularisation de Sobolev spatialement dépendante, outil d'inversion développé pour la première fois, a pu montrer son efficacité et sa souplesse d'utilisation, spécialement dans le cas de données fortement bruitées (10 dB).

L'algorithme de BFGS a ensuite été étendu au cas de l'approximation diffuse tridimensionnelle dans le régime fréquentiel ainsi qu'à l'équation de transfert radiatif bidimensionnelle dans le régime stationnaire. Concernant l'approximation diffuse, nous avons observé des meilleures performances de la régularisation de Sobolev uniforme dans le cas tridimensionnel par rapport au cas bidimensionnel (voir section 6.3). D'autre part, il a été constaté dans le chapitre 5 que les conclusions relatives aux performances de la régularisation de Sobolev spatialement dépendante, obtenues dans le cas de l'approximation diffuse bidimensionnelle, restent valables aux cas de l'approximation diffuse tridimensionnelle et de l'équation de transfert radiatif bidimensionnelle. De plus, il a été observé que les gradients de Sobolev spatialement dépendants peuvent empêcher l'établissement du phénomène dit de *crosstalk* entre les paramètres dans le cas de l'inversion de l'approximation diffuse tridimensionnelle, permettant une réduction des temps de calcul en plus de fournir de meilleures reconstructions.

Troisièmement, une méthode inverse multi-échelle basée sur les ondelettes associée à l'algorithme de BFGS a été développée pour la première fois (voir chapitre 4). Cette méthode, qui s'appuie sur une reformulation du problème inverse original en une suite de sous-problèmes inverses de la plus grande échelle à la plus petite, à l'aide de la transformée en ondelettes, a permis de faire face à la propriété de convergence locale de l'optimiseur et à la présence de nombreux minima locaux dans la fonction coût. Les résultats numériques ont montré que la méthode proposée est plus stable vis-à-vis de l'estimation initiale des propriétés radiatives et fournit des reconstructions finales plus précises que l'algorithme de BFGS ordinaire tout en nécessitant des temps de calcul semblables.

Cette méthode inverse multi-échelle, développée pour le modèle des sources diffuses dans le chapitre 4, a été étendue au modèle des sources collimatées dans la section 6.1. Nous avons observé que les conclusions du chapitre 4 restent valables pour le cas des sources collimatées et même, les améliorations de la méthode inverse multi-échelle par rapport à la méthode de BFGS ordinaire se sont trouvées plus prononcées. Le phénomène de crosstalk entre les paramètres, réputé difficile dans la communauté de la tomographie optique, a ensuite été étudié dans la section 6.2. Nous avons observé que la méthode multi-échelle permet d'atténuer ce phénomène par rapport à l'algorithme de BFGS ordinaire. Cependant, ces améliorations ne sont pas totalement satisfaisantes et d'autres méthodes devront être développées pour résoudre ce problème difficile.

## CONCLUSION ET PERSPECTIVES

### Conclusion générale

Dans cette thèse, nous nous sommes intéressés à l'étude du problème inverse de tomographie optique diffuse dans le régime fréquentiel.

Premièrement, une comparaison des algorithmes de Gauss-Newton amorti et de Broyden-Fletcher-Goldfarb-Shanno (BFGS) a été effectuée dans le cas bidimensionnel. Deux méthodes de régularisation ont été combinées pour chacun des deux algorithmes. Concernant l'algorithme de Gauss-Newton amorti, la réduction de la dimension de l'espace de contrôle et la pénalisation de Tikhonov ont été utilisées. Pour la méthode de BFGS, la réduction de la dimension de l'espace de contrôle et l'utilisation des gradients de Sobolev lors de l'extraction du gradient de la fonction coût sont, pour la première fois, simultanément appliquées. De plus, une méthode des gradients de Sobolev spatialement dépendants est pour la première fois développée dans le champ des problèmes inverses.

Les résultats obtenus ont permis d'aboutir aux conclusions suivantes :

- Le paramètre de régularisation quasi-optimal de Tikhonov, déterminé par la méthode « L-curve », est quasi-indépendant de la dimension de l'espace de contrôle dans le cas sous-déterminé ;
- La réduction de l'espace de contrôle, jusqu'à une certaine limite, conduit à des reconstructions finales plus précises pour les deux algorithmes ;
- La stratégie d'adimensionnalisation des variables de contrôle fournit des meilleurs résultats qu'une mise à l'échelle des gradients de la fonction de coût pour l'algorithme de BFGS ;
- Certaines valeurs du paramètre de régularisation pour la méthode de Sobolev uniforme permettent d'accéder à des reconstructions plus précises pour les coefficients d'absorption et de diffusion, excepté dans le cas très bruité (10 dB) ;

- La régularisation de Sobolev spatialement dépendante permet d'obtenir de meilleures reconstructions, notamment dans le cas très bruité, et ce, pour une large gamme de valeurs du paramètre de régularisation ;
- L'algorithme de BFGS surpasse celui de Gauss-Newton amorti en ce qui concerne la qualité des reconstructions obtenues, le temps de calcul ou encore la facilité de sélection du paramètre de régularisation.

Deuxièmement, faisant suite à la première conclusion énoncée ci-dessus, une étude plus théorique et plus poussée sur la quasi-indépendance du paramètre de pénalisation de Tikhonov optimal par rapport à la dimension de l'espace de contrôle dans les problèmes inverses d'estimation de fonctions spatialement dépendantes a été menée. Cette quasi-indépendance a été démontrée théoriquement puis vérifiée numériquement à la fois sur un problème inverse linéaire de conduction de la chaleur et sur le problème inverse non-linéaire de tomographie optique diffuse. Cette étude a conduit à la conclusion que le paramètre de Tikhonov pourra être déterminé en utilisant une paramétrisation des variables de contrôle associée à un maillage lâche afin de réduire les temps de calcul.

Troisièmement, une méthode inverse multi-échelle basée sur les ondelettes associée à l'algorithme de BFGS a été développée pour la première fois. La méthode a été appliquée pour le modèle des sources diffuses et celui des sources collimatées. Les ondelettes de Haar et de Daubechies D4 et D6 ont été utilisées. Une comparaison entre les méthodes de BFGS ordinaire et multi-échelle a été menée et a permis d'aboutir aux conclusions suivantes :

- La méthode multi-échelle basée sur les ondelettes de Haar fournit des meilleurs résultats que celle basée sur les ondelettes de Daubechies ;
- La méthode multi-échelle utilisant les ondelettes de Haar est plus stable vis-à-vis de l'estimation initiale des propriétés radiatives et fournit des reconstructions finales plus précises que l'algorithme de BFGS ordinaire tout en nécessitant des temps de calcul semblables ;
- L'amélioration de la qualité des reconstructions par la méthode multi-échelle par rapport à la méthode de BFGS ordinaire est plus importante dans le cas du modèle de la source collimatée que dans celui de la source diffuse.

Enfin, la résolution du problème inverse de tomographie optique diffuse par la méthode de BFGS ordinaire, combinant les méthodes de régularisation de la réduction de la dimension de l'espace de contrôle et des gradients de Sobolev, a été étendue au cas de l'approximation diffuse tridimensionnelle dans le régime fréquentiel et de l'équation de transfert radiatif bidimensionnelle dans le régime stationnaire. Les conclusions établies précédemment restent valables pour ces deux nouvelles modélisations.

### **Perspectives**

Loin de clore le sujet, les travaux menés au cours de cette thèse laissent entrevoir de nombreuses perspectives.

Concernant la méthode de régularisation des gradients de Sobolev spatialement dépendants, la recherche d'autres filtres plus performants que celui présenté dans cette étude pourrait être envisagée. Par ailleurs, cette régularisation a été couplée à la méthode multi-échelle et une amélioration des reconstructions a été observée. Cet effet cumulatif des deux régularisations devrait faire l'objet d'une prochaine publication.

Sur la quasi-indépendance du paramètre de pénalisation de Tikhonov optimal par rapport à la dimension de l'espace de contrôle dans les problèmes inverses, de nombreux résultats théoriques restent à étudier dans le cas non-linéaire. Par exemple, une extension à d'autres distributions de probabilité que la loi normale pour les fonctions radiatives à reconstruire serait intéressante.

En ce qui concerne la méthode inverse multi-échelle développée dans cette thèse, de nombreuses investigations restent à faire. Sur son application au cas du modèle des sources diffuses, la stratégie adoptée dans le chapitre 4 consistant à effectuer des projections unidimensionnelles sur bases d'ondelettes des mesures, des prédictions et du terme source de l'état adjoint lors du calcul de la fonction de coût et de ses gradients n'est peut-être pas optimale. La stratégie, plus simple, adoptée pour le cas des sources collimatées pourrait être considérée pour ce modèle et permettre, en plus de réduire les temps de calcul, d'obtenir de meilleurs résultats. Jusqu'à

présent, nous n'avons utilisé que la paramétrisation  $\mathcal{P}_1$  pour la méthode multi-échelle. Une étude sur les résultats obtenus en fonction de la régularité de la paramétrisation éléments finis utilisée pour les variables d'état, adjointe et de contrôle pourrait être menée. Des liens entre la régularité de la paramétrisation et la régularité de l'ondelette utilisée pourrait être établie. De plus, l'utilisation de transformées en ondelettes n'exhibant pas d'effets de bord indésirables, pour les autres ondelettes que celles de Haar, pourrait permettre une forte amélioration des résultats obtenus avec celles-ci. Une extension à des domaines tridimensionnels parallélépipédiques est à effectuer et ne devrait pas poser de difficulté particulière. En revanche, l'extension de la méthode à des géométries de formes quelconque devrait demander un certain effort. Deux approches pourraient être considérées pour résoudre ce problème. La première consisterait à immerger le domaine sondé dans un rectangle ou un parallélépipède de sorte à pouvoir effectuer les transformées en ondelettes. Ensuite, les valeurs des fonctions aux nœuds à l'extérieur du domaine devraient être fixées à 0 et les transformées en ondelettes devraient être corrigées de façon à prendre en compte ces valeurs nulles fictives. La seconde approche ferait appel à la topologie algébrique et consisterait à appliquer une déformation géométrique inversible permettant de passer de la géométrie du domaine sondé à une géométrie où les transformées en ondelettes sont possibles.

Les outils numériques développés dans le cadre de cette thèse sont aussi suffisamment génériques pour être utilisés dans la résolution d'autres problèmes inverses. Par exemple, la structure algorithmique, l'algorithme de minimisation de BFGS ainsi que les algorithmes de recherche linéaire ont récemment été réutilisés pour la résolution de problèmes inverses de tomographie optique basés sur l'ÉTR (Favennec *et al.*, 2015). D'autres outils développés tels que les gradients de Sobolev spatialement dépendants ou la méthode multi-échelle basée sur les ondelettes ne devraient pas, *a priori*, poser de sérieuses difficultés. Ces outils pourraient également être appliqués à d'autres problèmes inverses tels que ceux rencontrés en tomographie optique diffuse de fluorescence ou de bioluminescence, en tomographie d'impédance électrique ou encore en propagation d'ondes acoustiques, par exemple.

Enfin, une application en cours des travaux réalisés dans cette thèse concerne le développement d'une méthode de fonctionnalisation des matériaux cellulaires, dont un schéma du principe est donné dans la Figure 7.1. Premièrement, un matériaux cellulaire (ici une mousse réfractaire de carbure de silicium utilisée dans les centrales solaires thermiques) est généré numériquement à l'aide du logiciel iMorph (Rousseau *et al.*, 2012, 2013) (en bas à gauche). Les transferts radiatifs sont ensuite calculés dans cette mousse virtuelle par un code de lancer de rayon. Deuxièmement, les données issues de ce calcul sur les 6 faces du cube sont modélisées sous la forme d'une discrétisation éléments finis dans un milieu homogénéisé et vont constituer le terme source et les mesures pseudo-expérimentales (en bas à droite). Troisièmement, un algorithme de tomographie optique diffuse effectue la caractérisation des distributions spatiales des propriétés radiatives du milieu homogénéisé (en haut). Enfin, si les distributions des propriétés radiatives ne sont pas satisfaisantes du point de vue de l'application visée, la morphologie ou les propriétés des constituants de la mousse initiale sont mises à jour, et le processus est répété jusqu'à obtention des propriétés désirées. Cette méthode pourrait ainsi permettre le design de nouveaux matériaux à partir du seul outil numérique. Par ailleurs, l'extension de la méthode inverse multi-échelle aux domaines tridimensionnels parallélépipédiques apparaît, dans ce contexte, être très bien adaptée et pourrait ainsi contribuer au succès de la méthode.

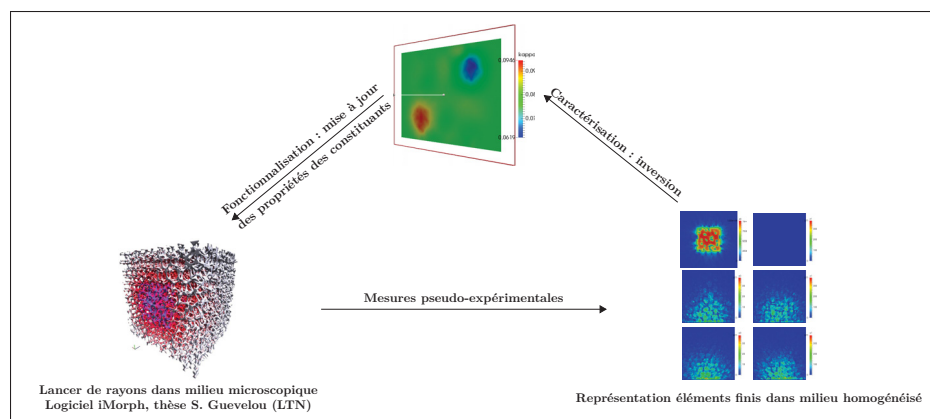


FIG. 7.1 Schématisation du principe de la méthode de fonctionnalisation des propriétés radiatives des matériaux.





## ANNEXE I

### CALCULATION OF THE COST FUNCTION GRADIENT WITH THE ADJOINT-STATE METHOD

The optimal method for large-scale optimization problems of the adjoint-state method is used in this appendix to calculate the cost function gradient. In order not to complicate the developments which follow, the experiment has been assumed to involve a single light source ( $K=1$ ) : the subscript  $l$  is omitted in the photon density notation while the strength and location of the collimated source is noted  $q_1$  and  $\mathbf{r}_1$ , respectively. Moreover, for  $f, g : \Omega \rightarrow \mathbb{R}$ , the following inner products are defined :

$$(f, g)_{L^2(\partial\Omega)} := \int_{\partial\Omega} fg \, d\zeta \quad (\text{A I-1})$$

$$(f, g)_{L^2(\partial\Omega_d)} := \int_{\partial\Omega} fg \mathbb{1}_{[\zeta \in \partial\Omega_d]}(\zeta) \, d\zeta = \sum_{j=1}^D f(\hat{\mathbf{r}}_j) g(\hat{\mathbf{r}}_j) \quad (\text{A I-2})$$

Let us first introduce the adjoint-state variable  $\varphi^* = \varphi^{r*} + i\varphi^{i*}$  to the state one  $\varphi = \varphi^r + i\varphi^i$ . The equation (2.1) implies that  $\varphi^r$  and  $\varphi^i$  are respectively solutions of the following partial differential equations :

$$-\nabla \cdot (D\nabla \varphi^r) + \kappa \varphi^r - \frac{2\pi\nu}{c} \varphi^i = q_1 \delta(\mathbf{r} - \mathbf{r}_1) \quad \forall \mathbf{r} \in \Omega \quad (\text{A I-3})$$

$$\varphi^r + \frac{A}{2\gamma} D\nabla \varphi^r \cdot n = 0 \quad \forall \zeta \in \partial\Omega \quad (\text{A I-4})$$

$$-\nabla \cdot (D\nabla \varphi^i) + \kappa \varphi^i + \frac{2\pi\nu}{c} \varphi^r = 0 \quad \forall \mathbf{r} \in \Omega \quad (\text{A I-5})$$

$$\varphi^i + \frac{A}{2\gamma} D\nabla \varphi^i \cdot n = 0 \quad \forall \zeta \in \partial\Omega \quad (\text{A I-6})$$

Remembering that  $\kappa(\varkappa) = \kappa_{ap}\varkappa$ ,  $\sigma(\varsigma) = \sigma_{ap}\varsigma$ ,  $D = (n_\Omega(\kappa + \sigma))^{-1}$  and introducing the residues of the forward models

$$\mathcal{R}(\kappa, \sigma, \varphi^r, \varphi^i) = -\nabla \cdot (D\nabla \varphi^r) + \kappa\varphi^r - \frac{2\pi\nu}{c}\varphi^i - q_1\delta(\mathbf{r} - \mathbf{r}_1) \quad (\text{A I-7})$$

$$\mathcal{S}(\kappa, \sigma, \varphi^r, \varphi^i) = -\nabla \cdot (D\nabla \varphi^i) + \kappa\varphi^i + \frac{2\pi\nu}{c}\varphi^r \quad (\text{A I-8})$$

and the Lagrange function

$$\mathcal{L}(\varkappa, \varsigma, \varphi^r, \varphi^i, \varphi^{r*}, \varphi^{i*}) = \mathcal{J}(\varphi^r, \varphi^i) + (\mathcal{R}, \varphi^{r*})_{L^2(\Omega)} + (\mathcal{S}, \varphi^{i*})_{L^2(\Omega)} \quad (\text{A I-9})$$

it can be observed that  $\mathcal{L}(\varkappa, \varsigma, \varphi^r, \varphi^i, \varphi^{r*}, \varphi^{i*}) = \mathcal{J}(\varphi^r, \varphi^i) = j(\varkappa, \varsigma)$  when  $\varphi^r$  and  $\varphi^i$  are solutions of eq. (A I-3) and eq. (A I-5), respectively. One can then deduces that for all  $\delta\varkappa \in L^2(\Omega)$  and for all  $\varphi^{r*}, \varphi^{i*} \in H^1(\Omega)$  :

$$j'(\gamma, \delta\varkappa) := (\nabla^\varkappa j, \delta\varkappa)_{L^2(\Omega)} \quad (\text{A I-10})$$

$$= \frac{\partial \mathcal{L}}{\partial \varkappa} \delta\varkappa + \frac{\partial \mathcal{L}}{\partial \varphi^r} \frac{\partial \varphi^r}{\partial \varkappa} \delta\varkappa + \frac{\partial \mathcal{L}}{\partial \varphi^i} \frac{\partial \varphi^i}{\partial \varkappa} \delta\varkappa \quad (\text{A I-11})$$

where the directional derivative  $j'(\varkappa, \delta\varkappa)$  is related to the partial differential sum of  $\mathcal{L}(\varkappa, \varsigma, \varphi^r, \varphi^i, \varphi^{r*}, \varphi^{i*})$ .

Let us note  $\varphi'_r = \frac{\partial \varphi^r}{\partial \varkappa} \delta\varkappa$  and  $\varphi'_i = \frac{\partial \varphi^i}{\partial \varkappa} \delta\varkappa$ . Eq. (A I-11) can be calculated by the sum of the following eight terms :

$$\left( \frac{\partial \mathcal{J}}{\partial \varphi^r}, \varphi'_r \right)_{L^2(\partial\Omega_d)} = \left( \frac{\varphi^r - \check{\varphi}^r}{|\check{\varphi}|^2}, \varphi'_r \right)_{L^2(\partial\Omega_d)} \quad (\text{A I-12})$$

$$\left( \frac{\partial \mathcal{J}}{\partial \varphi^i}, \varphi'_i \right)_{L^2(\partial\Omega_d)} = \left( \frac{\varphi^i - \check{\varphi}^i}{|\check{\varphi}|^2}, \varphi'_i \right)_{L^2(\partial\Omega_d)} \quad (\text{A I-13})$$

$$\left( \frac{\partial \mathcal{R}}{\partial \varphi^r}, \varphi'_r, \varphi^{r*} \right)_{L^2(\Omega)} = (-\nabla \cdot (D\nabla \varphi'_r) + \kappa\varphi'_r, \varphi^{r*})_{L^2(\Omega)} \quad (\text{A I-14})$$

$$\left( \frac{\partial \mathcal{R}}{\partial \varphi^i}, \varphi'_i, \varphi^{r*} \right)_{L^2(\Omega)} = \left( -\frac{2\pi\nu}{c}\varphi'_i, \varphi^{r*} \right)_{L^2(\Omega)} \quad (\text{A I-15})$$

and

$$\left( \frac{\partial \mathcal{S}}{\partial \varphi^r} \varphi'_r, \varphi^{i*} \right)_{L^2(\Omega)} = \left( \frac{2\pi v}{c} \varphi'_r, \varphi^{i*} \right)_{L^2(\Omega)} \quad (\text{A I-16})$$

$$\left( \frac{\partial \mathcal{S}}{\partial \varphi^i} \varphi'_i, \varphi^{i*} \right)_{L^2(\Omega)} = (-\nabla \cdot (D\nabla \varphi'_i) + \kappa \varphi'_i, \varphi^{i*})_{L^2(\Omega)} \quad (\text{A I-17})$$

$$\left( \frac{\partial \mathcal{R}}{\partial \varkappa} \delta \varkappa, \varphi^{r*} \right)_{L^2(\Omega)} = (\nabla \cdot (n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^r) + \kappa_{ap} \varphi^r \delta \varkappa, \varphi^{r*})_{L^2(\Omega)} \quad (\text{A I-18})$$

$$\left( \frac{\partial \mathcal{S}}{\partial \varkappa} \delta \varkappa, \varphi^{i*} \right)_{L^2(\Omega)} = (\nabla \cdot (n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^i) + \kappa_{ap} \varphi^i \delta \varkappa, \varphi^{i*})_{L^2(\Omega)} \quad (\text{A I-19})$$

Applying the Green identity twice to eqs. (A I-14) and (A I-17) and once to eqs. (A I-18) and (A I-19) yields :

$$\begin{aligned} \left( \frac{\partial \mathcal{R}}{\partial \varphi^r} \varphi'_r, \varphi^{r*} \right)_{L^2(\Omega)} &= (-\nabla \cdot (D\nabla \varphi^{r*}) + \kappa \varphi^{r*}, \varphi'_r)_{L^2(\Omega)} + (D\nabla \varphi^{r*} \cdot n, \varphi'_r)_{L^2(\partial\Omega)} \\ &\quad - (D\nabla \varphi'_r \cdot n, \varphi^{r*})_{L^2(\partial\Omega)} \end{aligned} \quad (\text{A I-20})$$

$$\begin{aligned} \left( \frac{\partial \mathcal{S}}{\partial \varphi^i} \varphi'_i, \varphi^{i*} \right)_{L^2(\Omega)} &= (-\nabla \cdot (D\nabla \varphi^{i*}) + \kappa \varphi^{i*}, \varphi'_i)_{L^2(\Omega)} + (D\nabla \varphi^{i*} \cdot n, \varphi'_i)_{L^2(\partial\Omega)} \\ &\quad - (D\nabla \varphi'_i \cdot n, \varphi^{i*})_{L^2(\partial\Omega)} \end{aligned} \quad (\text{A I-21})$$

$$\begin{aligned} \left( \frac{\partial \mathcal{R}}{\partial \varkappa} \delta \varkappa, \varphi^{r*} \right)_{L^2(\Omega)} &= - (n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^r, \nabla \varphi^{r*})_{L^2(\Omega)} + (\kappa_{ap} \varphi^r \delta \varkappa, \varphi^{r*})_{L^2(\Omega)} \\ &\quad + (n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^r \cdot n, \varphi^{r*})_{L^2(\partial\Omega)} \end{aligned} \quad (\text{A I-22})$$

$$\begin{aligned} \left( \frac{\partial \mathcal{S}}{\partial \varkappa} \delta \varkappa, \varphi^{i*} \right)_{L^2(\Omega)} &= - (n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^i, \nabla \varphi^{i*})_{L^2(\Omega)} + (\kappa_{ap} \varphi^i \delta \varkappa, \varphi^{i*})_{L^2(\Omega)} \\ &\quad + (n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^i \cdot n, \varphi^{i*})_{L^2(\partial\Omega)} \end{aligned} \quad (\text{A I-23})$$

Next, the directional derivatives of the Robin boundary conditions for eqs. (A I-3) and (A I-5) give :

$$-D\nabla \varphi'_r \cdot n = \frac{2\gamma}{A} \varphi'_r - n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^r \cdot n \quad (\text{A I-24})$$

and

$$-D\nabla\varphi'_i \cdot n = \frac{2\gamma}{A}\varphi'_i - n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^i \cdot n \quad (\text{A I-25})$$

Replacing the last terms of eqs. (A I-20) and (A I-21) by the second members of eqs. (A I-24) and (A I-25), the sum of eqs. (A I-20) to (A I-23),  $\Sigma_{\text{Temp}}$ , is equal to :

$$\begin{aligned} \Sigma_{\text{Temp}} = & \left( -\nabla \cdot (D\nabla \varphi^{r*}) + \kappa \varphi^{r*}, \varphi'_r \right)_{L^2(\Omega)} + \left( D\nabla \varphi^{r*} \cdot n, \varphi'_r \right)_{L^2(\partial\Omega)} \\ & + \left( -\nabla \cdot (D\nabla \varphi^{i*}) + \kappa \varphi^{i*}, \varphi'_i \right)_{L^2(\Omega)} + \left( D\nabla \varphi^{i*} \cdot n, \varphi'_i \right)_{L^2(\partial\Omega)} \\ & - \left( n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^r, \nabla \varphi^{r*} \right)_{L^2(\Omega)} + \left( \kappa_{ap} \varphi^r \delta \varkappa, \varphi^{r*} \right)_{L^2(\Omega)} \\ & - \left( n_\Omega \kappa_{ap} D^2 \delta \varkappa \nabla \varphi^i, \nabla \varphi^{i*} \right)_{L^2(\Omega)} + \left( \kappa_{ap} \varphi^i \delta \varkappa, \varphi^{i*} \right)_{L^2(\Omega)} \\ & + \left( \frac{2\gamma}{A} \varphi^{r*}, \varphi'_r \right)_{L^2(\partial\Omega)} + \left( \frac{2\gamma}{A} \varphi^{i*}, \varphi'_i \right)_{L^2(\partial\Omega)} \quad (\text{A I-26}) \end{aligned}$$

Thereby we obtain :

$$\begin{aligned} j'(\gamma; \delta \varkappa) & := (\nabla \varkappa j, \delta \varkappa)_{L^2(\Omega)} \\ & = \left( \kappa_{ap} (\varphi^r \varphi^{r*} + \varphi^i \varphi^{i*}) - n_\Omega \kappa_{ap} D^2 (\nabla \varphi^r \cdot \nabla \varphi^{r*} + \nabla \varphi^i \cdot \nabla \varphi^{i*}), \delta \varkappa \right)_{L^2(\Omega)} \\ & - \left( \nabla \cdot (D\nabla \varphi^{r*}) - \kappa \varphi^{r*} - \frac{2\pi \mathbf{v}}{c} \varphi^{i*}, \varphi'_r \right)_{L^2(\Omega)} \\ & - \left( \nabla \cdot (D\nabla \varphi^{i*}) - \kappa \varphi^{i*} + \frac{2\pi \mathbf{v}}{c} \varphi^{r*}, \varphi'_i \right)_{L^2(\Omega)} \\ & + \left( \frac{2\gamma}{A} \varphi^{r*} + D(\nabla \varphi^{r*} \cdot n), \varphi'_r \right)_{L^2(\partial\Omega)} + \left( \frac{2\gamma}{A} \varphi^{i*} + D(\nabla \varphi^{i*} \cdot n), \varphi'_i \right)_{L^2(\partial\Omega)} \\ & + \left( \frac{\varphi^r - \check{\varphi}^r}{|\check{\varphi}|^2}, \varphi'_r \right)_{L^2(\partial\Omega_d)} + \left( \frac{\varphi^i - \check{\varphi}^i}{|\check{\varphi}|^2}, \varphi'_i \right)_{L^2(\partial\Omega_d)} \quad (\text{A I-27}) \end{aligned}$$

Hence, solving the following adjoint equations for  $\varphi^* = \varphi^{r*} + \mathbf{i}\varphi^{i*}$ ,

$$-\nabla \cdot (D\nabla \varphi^*) + \left( \kappa - \frac{2\mathbf{i}\pi\nu}{c} \right) \varphi^* = 0 \quad \forall \mathbf{r} \in \Omega \quad (\text{A I-28})$$

$$\frac{2\gamma}{A} \varphi^* + D(\nabla \varphi^* \cdot \mathbf{n}) = -\frac{\varphi - \check{\varphi}}{|\check{\varphi}|^2} \mathbb{1}_{[\zeta \in \partial\Omega_d]}(\zeta) \quad \forall \zeta \in \partial\Omega, \quad (\text{A I-29})$$

it can be deduced that the cost function gradient related to  $\varkappa$  is :

$$(\nabla^\varkappa j, \delta \varkappa)_{L^2(\Omega)} = (\kappa_{ap}(\varphi^r \varphi^{r*} + \varphi^i \varphi^{i*}) - n_\Omega \kappa_{ap} D^2 (\nabla \varphi^r \cdot \nabla \varphi^{r*} + \nabla \varphi^i \cdot \nabla \varphi^{i*}), \delta \varkappa)_{L^2(\Omega)} \quad (\text{A I-30})$$

$$= (\text{Re} (\kappa_{ap} \varphi \bar{\varphi}^* - n_\Omega \kappa_{ap} D^2 \nabla \varphi \cdot \nabla \bar{\varphi}^*), \delta \varkappa)_{L^2(\Omega)} \quad (\text{A I-31})$$

An equivalent derivation for  $\varsigma$  would lead to :

$$(\nabla^\varsigma j, \delta \varsigma)_{L^2(\Omega)} = (-n_\Omega \sigma_{ap} D^2 (\nabla \varphi^r \cdot \nabla \varphi^{r*} + \nabla \varphi^i \cdot \nabla \varphi^{i*}), \delta \varsigma)_{L^2(\Omega)} \quad (\text{A I-32})$$

$$= (\text{Re} (-n_\Omega \sigma_{ap} D^2 \nabla \varphi \cdot \nabla \bar{\varphi}^*), \delta \varsigma)_{L^2(\Omega)} \quad (\text{A I-33})$$



## ANNEXE II

### PARAMÉTRISATION DES VARIABLES DE CONTRÔLE ET MÉTHODES DE RECHERCHE LINÉAIRE DANS UN CODE D'INVERSION DE L'APPROXIMATION DE DIFFUSION BASÉ SUR LE L-BFGS

Fabien Dubot<sup>1,2</sup>, Yann Favennec<sup>2</sup>, Benoit Rousseau<sup>2</sup>, Daniel Rousse<sup>1</sup>

<sup>1</sup> Chaire de recherche industrielle t3e, École de technologie supérieure,

1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

<sup>2</sup> Laboratoire de thermocinétique de Nantes, UMR CNRS 6607,

Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France

Comptes-Rendus Congrès de la Société Française de Thermique (SFT), Lyon, France, Juin 2014.

#### Résumé

La tomographie optique diffuse est une méthode inverse consistant à reconstruire les distributions spatiales des propriétés radiatives d'un milieu participant à partir de mesures simulées d'intensités lumineuses en frontière avec, pour modèle direct, l'équation de l'approximation de diffusion. Ce problème inverse est résolu par la minimisation d'une fonction de coût effectuée par l'algorithme de L-BFGS. Quatre types de paramétrisation éléments finis sont envisagés pour les cartes d'absorption et de diffusion. De plus, trois méthodes de recherche linéaire sont considérées dans l'optimiseur : les méthodes de la section d'or, de l'interpolation quadratique et une méthode de « recherche inexacte ». Des résultats numériques bidimensionnels viennent illustrer cette étude.

#### 2. Introduction

La modélisation des transferts radiatifs dans un milieu semi-transparent nécessite, entre autres, une connaissance fine des propriétés physiques de ce milieu. Parmi ces données, une attention toute particulière doit être portée à l'estimation des fonctions d'absorption et de diffusion. En effet, ces deux coefficients (souvent fonctions de l'espace) sont directement impliqués dans les

équations qui décrivent la physique des transferts : leur connaissance précise constitue donc l'un des ingrédients pour une simulation précise de la physique considérée. Ces propriétés radiatives peuvent être estimées de façon dite directe par une connaissance précise de la géométrie et de la morphologie à échelle micro-métrique (par exemple via la  $\mu$ -tomographie X) et la composition du ou des matériaux constitutants (Rousseau *et al.*, 2007). Lorsque certaines de ces données sont manquantes, ou du moins imprécises, il convient d'utiliser des approches alternatives ; l'inversion de données en fait partie. Dans l'étude présentée ici, l'inversion de mesure consiste à « trouver » un jeu de fonctions d'absorption  $\kappa(\mathbf{r})$  et de diffusion réduit  $\sigma(\mathbf{r})$  tel que les prédictions issues de l'intégration des équations aux dérivées partielles soient aussi proches que possible des mesures associées. La minimisation de cet écart (qui est en fait une erreur) constitue le problème d'optimisation. Comme la dimension des paramètres recherchés peut être relativement grande, les algorithmes utilisés reposent nécessairement sur les méthodes de gradient. Parmi elles, les méthodes de BFGS (Broyden-Fletcher-Goldfarb-Shanno) et les versions associées à mémoire limitées (L-BFGS) ont déjà fait leur preuve en tomographie optique (TO) (Balima *et al.*, 2013; Dubot *et al.*, 2015b). Un nombre impressionnant de versions et sous-versions de ces algorithmes est présenté dans la littérature. Par exemple, certaines recherches linéaires peuvent être très efficaces pour un type de problème d'optimisation donné mais peuvent être inadaptées ou peu efficaces ou moins bons qu'un autre pour le problème spécifique de reconstructions de cartes de  $\kappa(\mathbf{r}) - \sigma(\mathbf{r})$ . C'est la raison pour laquelle différents outils de recherche linéaire ont été comparés dans le cadre spécifique de cette étude. Dans le même esprit, différents types de paramétrisation de Lagrange ont été comparés, autant d'un point de vue temps de calcul que de la qualité des résultats. Le modèle adopté pour cette étude est celui approché dit de l'approximation diffuse en bidimensionnel.

### 3. Formulation du problème inverse

Le problème inverse de TO consiste à retrouver les évolutions spatiales des propriétés radiatives constitutives d'un matériau semi-transparent. Dans le cadre de cet article, le modèle de propagation de la lumière est basé sur l'approximation de diffusion (AD). Ce modèle approché décrit de façon satisfaisante le modèle plus complexe – mais bien plus difficile à résoudre –



basé sur l'Équation de Transfert Radiatif (ÉTR), lorsque le matériau considéré est absorbant fortement diffusant, c'est-à-dire lorsque  $0 \ll \kappa \ll \sigma$ , et lorsque la taille du domaine est beaucoup plus grande que la longueur de transport. L'étude détaillée faisant référence sur le passage de l'ÉTR à l'AD a été faite relativement récemment par Arridge (Arridge, 1999). D'autre part, le régime fréquentiel est considéré pour cette étude : d'un côté le régime stationnaire est facile à résoudre mais donne trop peu d'information pour l'inversion et, à l'opposé, le régime temporel donne beaucoup d'information mais est trop coûteux en temps de calcul ; le domaine fréquentiel constitue donc une approche intermédiaire pour laquelle les modèles mathématiques sont faciles à résoudre et les données sont considérées comme suffisantes pour la résolution du problème inverse (Arridge, 1999).

Dans le contexte de cette étude, les paramètres suivants sont supposés connus : la géométrie bidimensionnelle  $\Omega$ , la fréquence de modulation  $\omega$ , les intensités des sources lumineuses  $f_\omega$  sur les emplacements  $\partial\Omega_s$  ainsi que divers paramètres usuels tels que  $\gamma_2$  et  $A$  impliqués dans les équations suivantes et qui se trouvent aisément dans la littérature.

Le problème dit *direct* se lit de la sorte : « *Connaissant les propriétés d'absorption  $\kappa$  et de diffusion  $\sigma$  spatio-dépendantes dans  $\Omega$ , déterminer l'intensité lumineuse au niveau des détecteurs  $\partial\Omega_d$  par la recherche de  $\varphi : \Omega \mapsto \mathbb{C}$  satisfaisant :*

$$\begin{aligned} -\nabla \cdot [(2(\kappa + \sigma))^{-1} \nabla \varphi] + \left[ \kappa + \frac{2\pi i \omega}{c} \right] \varphi &= 0, \quad \forall \mathbf{r} \in \Omega \\ \varphi + \frac{A}{2\gamma_2} (2(\kappa + \sigma))^{-1} \nabla \varphi \cdot \mathbf{n} &= \frac{f_\omega}{\gamma_2} \mathbb{1}_{[\zeta \in \partial\Omega_s]} \quad \forall \zeta \in \partial\Omega \end{aligned} \quad (\text{A II-1})$$

À l'opposé, le problème inverse de TO se lit ... « *Connaissant l'intensité lumineuse  $\varphi$  au niveau des détecteurs  $\partial\Omega_d$ , déterminer « au mieux » les évolutions spatiales des propriétés radiatives  $\kappa$  et  $\sigma$*  ». Bien sûr, la résolution du problème inverse est intimement liée à celle du problème direct basé sur (A II-1). De plus, contrairement au problème direct, ce type de problème inverse est réputé difficile car mal posé au sens de Hadamard (Hadamard, 1902). Différentes stratégies de régularisation sont donc potentiellement utilisées : utilisation de terme

pénalisant de type Tikhonov, réduction de l'espace de contrôle et utilisation des gradients de Sobolev, suivant (Balima *et al.*, 2013). Notons que seules les deux dernières techniques de régularisations énoncées seront employées dans cet article. Dans la suite, on suppose que l'on cherche  $\kappa \in \Lambda$  et  $\sigma \in \Lambda$ , avec  $\Lambda \subset L_2(\Omega)$ . La fonction de coût  $j : \Lambda^2 \mapsto ]0, \infty[$  est définie par :

$$j(\kappa, \sigma) := \mathcal{J}(\varphi) = \frac{1}{2} \sum_{k=1}^K \int_{\partial\Omega_d} \left| \frac{\varphi - \check{\varphi}}{\check{\varphi}} \right|^2 (k) \, d\mathbf{r} \quad (\text{A II-2})$$

où l'indice  $k$  désigne la source utilisée et, par extension,  $(\cdot)(k)$  désigne l'évaluation de  $(\cdot)$  relatif au test  $k$ . Le jeu de propriétés  $(\kappa, \sigma)$  est supposé retrouvé lorsque

$$j(\kappa, \sigma) \approx \varepsilon^2 \quad (\text{A II-3})$$

où  $\varepsilon^2$  est la variance de l'erreur de mesure intégrée et sommée. Dans la pratique, la condition (A II-3) n'est jamais atteinte pour les difficiles problèmes de tomographie ; le problème s'écrit plutôt :

$$\min_{(\kappa, \sigma) \in \Lambda^2} j(\kappa, \sigma) \quad (\text{A II-4})$$

#### 4. Optimiseur

La résolution du problème inverse est effectuée par la minimisation de la fonction de coût par l'algorithme d'optimisation de L-BFGS (Liu et Nocedal, 1989). Comme il a été montré dans (Dubot *et al.*, 2015b), cet optimiseur, associé à une méthode de recherche linéaire inexacte, est particulièrement efficace pour la résolution du problème de TO. Notamment, contrairement aux algorithmes d'inversion basés sur l'inversion de matrices tel l'algorithme de Gauss-Newton, l'algorithme de L-BFGS ne nécessite pas l'utilisation de méthodes de régularisation de type pénalisation de Tikhonov dont la détermination du paramètre de régularisation peut-être difficile à évaluer. Le principe des méthodes à direction de descente, l'algorithme de BFGS et sa routine de mise à jour de la direction suivant (Nocedal, 1980), sont brièvement rappelés ci-dessous. Puis, quelques méthodes de recherche linéaire qui seront testées numériquement par la suite sont exposées.

#### 4.1 L-BFGS

De manière générale les algorithmes d'optimisation à direction de descente nécessitent, d'une part, le calcul d'une direction de descente à chaque itération et, d'autre part, la détermination d'un pas de descente. Par définition,  $d \in \Lambda^2$  est une direction de descente de  $j$  en  $(\kappa, \sigma) \in \Lambda^2$  si

$$j'(\kappa, \sigma) \cdot d < 0 \quad (\text{A II-5})$$

Ainsi, en appliquant la définition de la dérivée, il vient que si  $d$  est une direction de descente :

$$j((\kappa, \sigma) + \alpha d) < j(\kappa, \sigma) \quad \text{pour tout } \alpha > 0 \text{ suffisamment petit.} \quad (\text{A II-6})$$

On observe alors que  $j$  décroît strictement dans la direction  $d$  et que la taille du déplacement le long de  $d$  est donnée par  $\alpha d$ . La direction de descente à l'itération  $m$  est définie par :

$$d_m = -H_m^{-1} g_m \quad (\text{A II-7})$$

où  $g_m$  est le gradient de la fonctionnelle à minimiser et  $H_m$  est une matrice définie positive. Par défaut (et dans cette étude) la matrice  $H_m^{-1}$  est initialisée à la matrice identité  $I_d$ . Sa formule de mise à jour est quant à elle donnée par :

$$H_{m+1}^{-1} = H_m^{-1} + \frac{s_m s_m^\top}{y_m^\top s_m} \left[ \frac{y_m^\top H_m^{-1} y_m}{y_m^\top s_m} + I_d \right] - \frac{1}{y_m^\top s_m} \left[ s_m y_m^\top H_m^{-1} + H_m^{-1} y_m s_m^\top \right] \quad (\text{A II-8})$$

où  $s_m = (\kappa_{m+1}, \sigma_{m+1}) - (\kappa_m, \sigma_m)$  et  $y_m = g_{m+1} - g_m$ . Suivant (Nocedal, 1980), la matrice  $H_m^{-1}$  n'est jamais calculée explicitement dans l'algorithme dans un souci d'efficacité. Seuls les vecteurs  $s_m$  et  $y_m$  sont stockés au fur et à mesure des itérations et permettent à eux seuls la mise à jour de la direction, éq. (A II-7).

Notons qu'au début de l'optimisation l'algorithme de BFGS se comporte comme celui de la plus grande pente ( $H_0^{-1} = I_d$ ). En revanche, au fur et à mesure des itérations et grâce à la

formule de mise à jour, éq. (A II-7), l'algorithme de BFGS est conçu pour tendre vers celui de Newton ( $H_m^{-1} \approx [\nabla^2 j]^{-1}$ ).

## 4.2 Méthodes de recherche linéaire

Les méthodes de recherche linéaire, qui ont pour but la détermination du pas de descente  $\alpha$ , éq. (A II-6), consistent en la résolution du problème de minimisation unidimensionnel suivant :

$$\alpha^* \approx \arg \min_{\alpha > 0} f(\alpha) \quad (\text{A II-9})$$

où  $f(\alpha) = j((\kappa, \sigma) + \alpha d)$ . Ces méthodes tentent de répondre à un double objectif : d'une part, le pas doit faire décroître  $j$  suffisamment et, d'autre part, il ne doit pas être trop petit sous peine de convergence des itérés vers un point non stationnaire. Or ces deux objectifs sont souvent contradictoires car l'éq. (A II-6) est généralement satisfaite pour des  $\alpha > 0$  arbitrairement petit.

Trois méthodes de recherche linéaire sont expérimentées dans l'optimiseur : (i) la méthode de la section d'or, (ii) la méthode de l'interpolation quadratique et (iii) une méthode de recherche inexacte due à Fletcher (Fletcher, 1980, p. 126).

### 4.2.1 Méthodes de réduction d'intervalle

Les méthodes de la section d'or et de l'interpolation quadratique sont toutes les deux des méthodes de réduction d'intervalle. Supposant  $\alpha^* \in [\alpha_L, \alpha_U]$ , leur principe repose sur une réduction répétée de cet intervalle initial, à l'aide d'évaluations de la fonction de coût en différents points prédéterminés par la méthode choisie et les bornes de cet intervalle et ce, jusqu'à ce que la longueur de l'intervalle soit plus petite qu'une taille prédéfinie.

Pour la méthode de la section d'or, le rapport des longueurs entre deux intervalles successifs est donné par le nombre d'or,  $\Xi = \frac{1+\sqrt{5}}{2}$ . Pour la méthode de l'interpolation quadratique, une approximation d'ordre deux de la fonction  $f$  est considérée. Il s'en suit que le point P minimisant cette fonction quadratique peut-être calculé analytiquement dans l'intervalle initial.

L'intervalle suivant est alors déterminé judicieusement à l'aide des évaluations de la fonction  $f$  aux bornes de l'intervalle et au point  $P$ .

L'implémentation de ces deux méthodes a été réalisée à partir des algorithmes 4.2 et 4.3 donnés dans (Antoniou et Lu, 2007). L'intervalle initial retenu pour les illustrations numériques est  $[0, 2]$  tandis que la recherche linéaire s'arrête lorsque la longueur de l'intervalle devient plus petite que 0.1.

#### 4.2.2 Méthodes inexactes

La philosophie des méthodes de recherche linéaire inexactes est différente dans le sens où, au lieu de résoudre l'éq. (A II-9), le pas  $\alpha$  est accepté dès que certaines conditions sont remplies. On a alors tout un intervalle où le pas est acceptable, ce qui rend la recherche linéaire plus aisée. De nos jours, les conditions à remplir par le pas considérées sont généralement les conditions de Wolfe, données par :

$$f(\alpha) \leq f(0) + c_1 \alpha f'(0) \quad (\text{A II-10})$$

$$f'(\alpha) \geq c_2 f'(0) \quad (\text{A II-11})$$

où les constantes  $c_1$  et  $c_2$  vérifient  $0 < c_1 < c_2 < 1$ . La première de ces conditions, aussi appelée condition d'Armijo, permet d'éviter que le pas  $\alpha$  retenu soit trop grand. La seconde quant à elle permet d'éviter que le pas  $\alpha$  retenu soit trop petit. Notons que plus  $c_2$  est petit, plus la recherche linéaire sera précise mais plus l'effort de calcul sera important.

La recherche linéaire inexacte considérée dans ce papier est due à Fletcher (Fletcher, 1980, p. 126). Son implémentation couplée au BFGS a été décrite dans l'algorithme 7.3 de (Antoniou et Lu, 2007). C'est cette dernière qui sera numériquement testée dans la suite.

### 5. Gradient de la fonction de coût

La méthode utilisée pour le calcul des composantes du gradient de la fonction de coût est bien entendu basée sur l'état adjoint puisque cette méthode allie rapidité et efficacité. Une démonstration complète pourra être trouvée dans (Dubot *et al.*, 2015b). Cette démonstration est basée sur la construction du lagrangien suivant (Favennec *et al.*, 2011). On y montre que lorsque le problème dit adjoint qui consiste à chercher  $\varphi^* : \Omega \mapsto \mathbb{C}$  est solution de :

$$\begin{aligned} -\nabla \cdot ((2(\kappa + \sigma))^{-1} \nabla \varphi^*) + (\kappa - \frac{2\pi i \omega}{c}) \varphi^* &= 0, & \forall \mathbf{r} \in \Omega \\ \frac{2\gamma_2}{A} \varphi^* + (2(\kappa + \sigma))^{-1} \nabla \varphi^* \cdot \mathbf{n} &= -\frac{1}{|\check{\varphi}|^2} (\varphi - \check{\varphi}) \mathbb{1}_{[\zeta \in \partial \mathcal{D}_d]}(\zeta) & \forall \zeta \in \partial \Omega \end{aligned} \quad (\text{A II-12})$$

et si le produit scalaire complexe  $L_2(\Omega)$  est utilisé lors de l'extraction du gradient, alors les composantes du gradient de la fonction de coût s'écrivent :

$$\begin{aligned} \nabla_{\kappa} j &= \sum_{k=1}^K \text{Re} (\varphi \bar{\varphi}^* - 2 \times (2(\kappa + \sigma))^{-2} \nabla \varphi \cdot \nabla \bar{\varphi}^*) \\ \nabla_{\sigma} j &= \sum_{k=1}^K \text{Re} (-2 \times (2(\kappa + \sigma))^{-2} \nabla \varphi \cdot \nabla \bar{\varphi}^*) \end{aligned} \quad (\text{A II-13})$$

## 6. Résultats numériques

### 6.1 Présentation du test

La reconstruction simultanée des cartes d'absorption et de diffusion réduit est effectuée sur un domaine bidimensionnel carré  $\Omega$  de côté 6 cm,  $\Omega = [-3, 3]^2$ . Les distributions spatiales des propriétés cibles, dont les valeurs sont représentatives de celles de tissus humains où la TO y est appliquée pour la détection de tumeurs, sont définies par :

$$\begin{aligned} \kappa(\mathbf{r}) &= 0.06 \text{ cm}^{-1}, \quad \sigma(\mathbf{r}) = 10 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega_1, & \Omega_1 &= [-1.8, -0.6]^2 \text{ cm}^2 \\ \kappa(\mathbf{r}) &= 0.1 \text{ cm}^{-1}, \quad \sigma(\mathbf{r}) = 30 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega_2, & \Omega_2 &= [0.6, 1.8]^2 \text{ cm}^2 \\ \kappa(\mathbf{r}) &= 0.08 \text{ cm}^{-1}, \quad \sigma(\mathbf{r}) = 20 \text{ cm}^{-1}, \quad \mathbf{r} \in \Omega \setminus (\Omega_1 \cup \Omega_2) \end{aligned} \quad (\text{A II-14})$$

Concernant l'acquisition des données synthétiques, 4 sources et 8 détecteurs carrés sont disposés autour du domaine  $\Omega$ , chacun d'une taille de 1,2 cm. Leur répartition, identique sur chaque côté du domaine, est donnée par :  $r_1$  (ou  $r_2$ )  $\in [-0.6, 0.6]$  cm pour les sources et

$r_1$  (ou  $r_2$ )  $\in [-2.4, -1.2] \cup [1.2, 2.4]$  cm pour les détecteurs. L'espace d'approximation des variables d'état et adjointe consiste en une paramétrisation éléments finis de type Lagrange  $\mathcal{P}_4$  basée sur un maillage régulier du domaine  $\Omega$  composé de 60 segments par côté. Signalons que cet espace d'approximation doit être suffisamment fourni afin de pouvoir reproduire les mesures expérimentales, mais ne doit pas trop l'être pour ne pas alourdir la résolution du problème inverse. Un maillage différent a été utilisé pour la création des données synthétiques (plus fin : composé de 65 segments par côté) dans le but d'éviter le « crime inverse » (Colton et Kress, 1992). De plus, un bruit blanc gaussien de 30 dB a été appliqué aux mesures pseudo-expérimentales aux noeuds des détecteurs afin de simuler « au mieux » des données réelles.

Notons enfin que le choix d'une paramétrisation  $\mathcal{P}_4$  comme espace d'approximation des données synthétiques et des variables d'état et adjointe s'appuie sur le fait qu'une comparaison entre les espaces d'approximation éléments finis  $\mathcal{P}_0$ ,  $\mathcal{P}_1$ ,  $\mathcal{P}_2$  et  $\mathcal{P}_3$  pour les variables de contrôle est effectuée ci-dessous. Ainsi, pour ne pas favoriser une reconstruction plutôt qu'une autre par corrélation entre espaces d'approximation, celui des données synthétiques et des variables d'état et adjointe a dû être choisi différent des éléments finis  $\mathcal{P}_0$ ,  $\mathcal{P}_1$ ,  $\mathcal{P}_2$  et  $\mathcal{P}_3$ , et s'est donc porté sur l'élément fini  $\mathcal{P}_4$ .

## 6.2 Comparaison des paramétrisations éléments finis

La comparaison entre les différents espaces d'approximation est faite à temps CPU fixé à 15 min et pour un maillage donné. La méthode de recherche linéaire utilisée est celle de Fletcher avec les paramètres  $\rho_F = 0.1$ ,  $\sigma_F = 0.7$  comme recommandé dans (Antoniou et Lu, 2007, p. 113). La Figure II-1 ainsi que le tableau II-1 montrent que l'espace élément fini  $\mathcal{P}_1$  est celui qui fournit les meilleurs résultats, confirmant les résultats de (Balima *et al.*, 2013).

## 6.3 Comparaison des méthodes de recherche linéaire

La comparaison entre les différentes méthodes de recherche linéaire est faite à temps CPU fixé à 20 min et pour une paramétrisation  $\mathcal{P}_1$  donnée. Les paramètres intervenant dans les

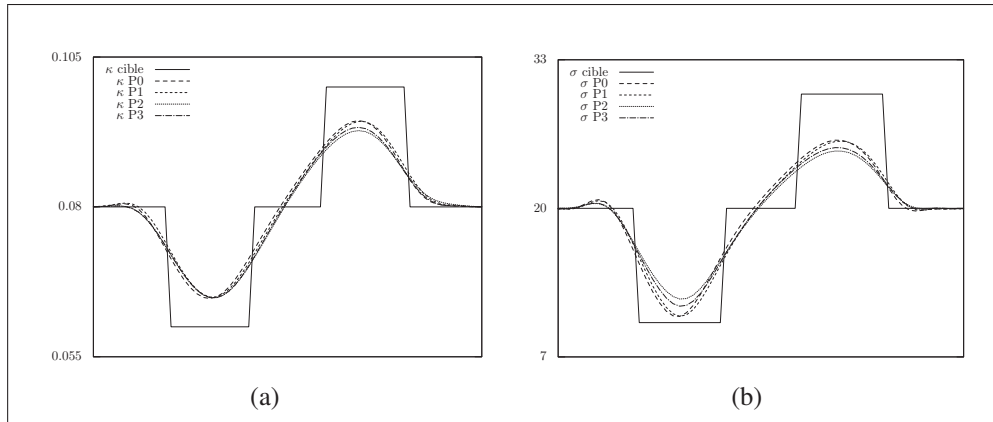


Figure-A II-1 Coupes 1d de  $\kappa$  et  $\sigma$ , respectivement en (a) et (b).

Tableau-A II-1 Erreurs  $L^2(\Omega)$  entre les reconstructions finales et les propriétés cibles.

	Erreur initiale	$\mathcal{P}_0$	$\mathcal{P}_1$	$\mathcal{P}_2$	$\mathcal{P}_3$
$\mathbf{e}_\kappa$	0,071	0,045	0,044	0,053	0,048
$\mathbf{e}_\sigma$	0,141	0,092	0,087	0,093	0,091

méthodes de la section d'or (GOL) et de l'interpolation quadratique (QUA) sont ceux donnés précédemment. Concernant la recherche linéaire de Fletcher, 3 jeux de paramètres sont testés : la version « précise » (PRE) avec  $\rho_F = 0.01$  et  $\sigma_F = 0.07$ , la version « imprécise » (IMP) avec  $\rho_F = 0.5$  et  $\sigma_F = 0.99$ , et une version « intermédiaire » (INT) avec  $\rho_F = 0.1$  et  $\sigma_F = 0.7$ .

La Figure II-2 ainsi que le tableau II-2 montrent que (i) la méthode de la section d'or fournit les moins bons résultats – ceci peut être justifié par le faible nombre d'itérations dans le temps imparti ; (ii) les trois versions de Fletcher fournissent des résultats équivalents – le faible nombre d'itérations effectué dans la version (PRE) est compensé par la grande précision demandée ; (iii) la méthode d'interpolation quadratique est celle qui fournit les meilleures reconstructions. Une reconstruction utilisant les meilleures données (paramétrisation  $\mathcal{P}_1$  couplée à la méthode quadratique de recherche linéaire) est présentée dans la Figure II-3. On peut remarquer que les deux inclusions sont relativement bien retrouvées comparativement à ce que l'on peut retrouver dans la littérature. On observe de plus que ces reconstructions, obtenues après un temps CPU



fixé à 30 min, sont très proches de celles présentées dans la Figure II-2, illustrant en quelque sorte la convergence de la méthode au bout de 20 min.

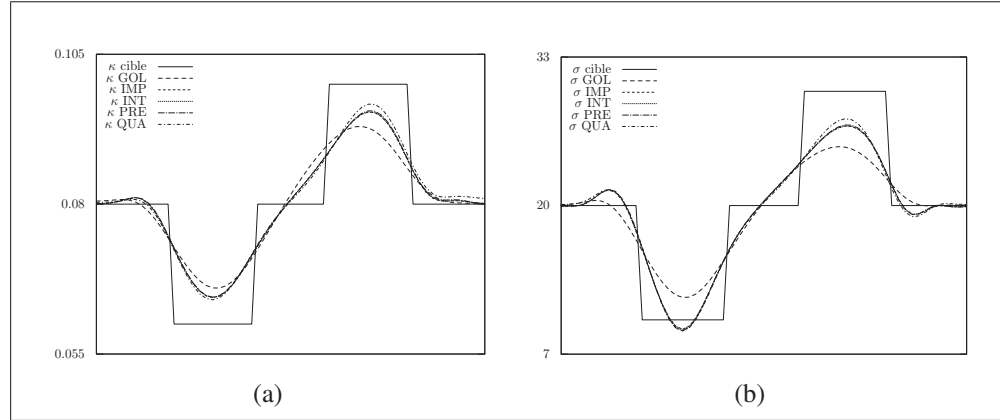


Figure-A II-2 Coupes 1d de  $\kappa$  et  $\sigma$ , respectivement en (a) et (b).

Tableau-A II-2 Erreurs  $L^2(\Omega)$  entre les reconstructions finales et les propriétés cibles.

	$e_\kappa$	$e_\sigma$	Itérations dans BFGS
<b>Erreur initiale</b>	0,0709	0,141	
<b>IMP</b>	0,0419	0,082	71
<b>INT</b>	0,0418	0,082	71
<b>PRE</b>	0,0419	0,082	38
<b>GOL</b>	0,0470	0,092	24
<b>QUA</b>	0,0412	0,078	68

## 7. Conclusion

Les résultats numériques présentés ont montré que la paramétrisation  $\mathcal{P}_1$  et la méthode de recherche linéaire de l'interpolation quadratique ont donné les meilleurs résultats. Il convient cependant de rester prudent notamment pour la méthode de l'interpolation quadratique qui nécessite l'utilisation d'un intervalle initial. A priori, ces conclusions demeurent valables pour l'ÉTR et/ou pour des géométries tridimensionnelles.

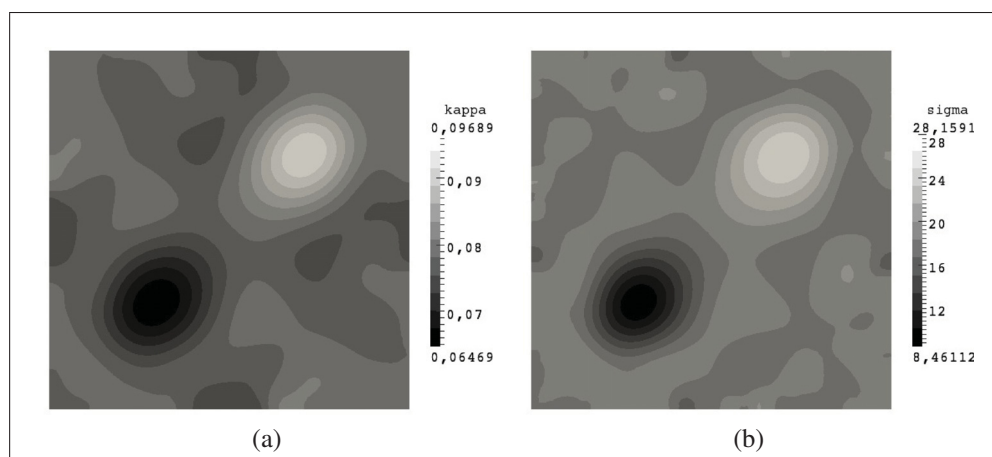


Figure-A II-3 Reconstructions  $\kappa$  et  $\sigma$  après 30 min, respectivement en (a) et (b). Paramétrisation  $\mathcal{P}_1$  et méthode de l'interpolation quadratique.

### ANNEXE III

#### NUMERICAL VALIDATION OF THE FORWARD MODEL SOLUTION

Transient Diffuse Approximation (DA) is given by :

$$\begin{cases} \frac{1}{c} \frac{\partial \varphi_s}{\partial t}(\mathbf{r}, t) - \nabla \cdot [D(\mathbf{r}) \nabla \varphi_s(\mathbf{r}, t)] + \kappa(\mathbf{r}) \varphi_s(\mathbf{r}, t) &= X_s(\mathbf{r}, t), \forall \mathbf{r} \in \Omega, \forall t > 0 \\ 2\gamma \varphi_s(\mathbf{r}, t) + AD(\mathbf{r}) \nabla \varphi_s(\mathbf{r}, t) \cdot \mathbf{n} &= 2x_s(\mathbf{r}, t), \forall \mathbf{r} \in \partial\Omega, \forall t > 0 \end{cases} \quad (\text{A III-1})$$

where  $X_s$  is an internal radiation source [ $\text{W cm}^{-3}$ ],  $x_s$  is the time-dependent signal associated with the  $s$ -th diffuse source [ $\text{W cm}^{-2}$ ] whose definition domain is denoted by  $\partial\Omega_s \subset \partial\Omega$ ,  $\varphi_s : (\Omega \times \mathbb{R}) \mapsto \mathbb{R}$  is the  $s$ -th photon density field [ $\text{W cm}^{-2}$ ],  $D(\mathbf{r}) = (n_\Omega [\kappa(\mathbf{r}) + \sigma(\mathbf{r})])^{-1}$  is the macroscopic scattering coefficient [ $\text{cm}$ ],  $n_\Omega$  ( $= 2, 3$ ) is the dimension of  $\Omega$ ,  $\kappa$  and  $\sigma = (1 - g)\sigma_s$  are the absorption and reduced scattering coefficients ( $\sigma_s$  is the scattering coefficient) [ $\text{cm}^{-1}$ ],  $g$  is the asymmetry factor which equals to the average cosine of the scattering angle,  $c$  is the speed of light in  $\Omega$  (constant) [ $\text{cm s}^{-1}$ ],  $\mathbf{n}$  is the unit normal vector to the boundary of  $\Omega$ ,  $\gamma$  is a constant parameter depending on  $n_\Omega$ , and  $A$  is a parameter which characterizes the reflection on the frontier  $\partial\Omega$ .

Fourier transforms are defined by :

$$\widehat{\varphi}(\mathbf{r}, f) = \int_{\mathbb{R}} \varphi(\mathbf{r}, t) \exp(-2i\pi f t) dt \quad \longleftrightarrow \quad \varphi(\mathbf{r}, t) = \int_{\mathbb{R}} \widehat{\varphi}(\mathbf{r}, f) \exp(2i\pi f t) df \quad (\text{A III-2})$$

The following relations

$$\widehat{\frac{\partial \varphi}{\partial t}}(\mathbf{r}, f) = 2i\pi f \widehat{\varphi}(\mathbf{r}, f) \quad \text{and} \quad \widehat{\nabla \varphi}(\mathbf{r}, f) = \nabla \widehat{\varphi}(\mathbf{r}, f) \quad (\text{A III-3})$$

are used to obtain the DA model at frequency  $f = f_0$  :

$$\begin{cases} -\nabla \cdot [D(\mathbf{r}) \nabla \widehat{\varphi}_s(\mathbf{r}, f_0)] + \left[ \kappa(\mathbf{r}) + \frac{2\pi i f_0}{c} \right] \widehat{\varphi}_s(\mathbf{r}, f_0) &= \widehat{X}_s(\mathbf{r}, f_0), \forall \mathbf{r} \in \Omega \\ 2\gamma \widehat{\varphi}_s(\mathbf{r}, f_0) + AD(\mathbf{r}) \nabla \widehat{\varphi}_s(\mathbf{r}, f_0) \cdot \mathbf{n} &= 2\widehat{x}_s(\mathbf{r}, f_0), \forall \mathbf{r} \in \partial\Omega \end{cases} \quad (\text{A III-4})$$

Suppose that the spectrum of the solution  $\widehat{\varphi}_s(\mathbf{r}, f)$  of eq. (A III-1) is composed of frequencies  $(f_i)_{i=1}^F$ . Then, the inverse Fourier transform provides the time-dependent solution if complex solutions of (A III-4) for each  $f_i, i = 1, \dots, F$  are known :

$$\varphi_s(\mathbf{r}, t) = \text{Re} \left[ \sum_{i=1}^F \exp(2i\pi f_i t) \widehat{\varphi}_s(\mathbf{r}, f_i) \right] \quad (\text{A III-5})$$

Then, suppose  $\widehat{\varphi}_s \in \widehat{H}^1(\Omega)$  and let  $v \in H^1(\Omega)$  be a test function. Then, the variational formulation of (A III-4) is given by :

$$a(\varphi_s, v) = l(v), \quad \forall v \in H^1(\Omega) \quad (\text{A III-6})$$

with

$$a(\varphi_s, v) = \int_{\Omega} D(\mathbf{r}) \nabla \widehat{\varphi}_s \cdot \nabla \bar{v} + \left( \kappa(\mathbf{r}) + \frac{2i\pi f}{c} \right) \widehat{\varphi}_s \bar{v} \, d\mathbf{r} + \int_{\partial\Omega} \frac{2\gamma}{A} \widehat{\varphi}_s \bar{v} \, d\mathbf{r} \quad (\text{A III-7})$$

and

$$l(v) = \int_{\Omega} \widehat{X}_s \bar{v} \, d\mathbf{r} + \int_{\partial\Omega} \frac{2}{A} \widehat{x}_s \bar{v} \, d\mathbf{r}. \quad (\text{A III-8})$$

First, the following source terms are considered :

$$\begin{cases} X_s(\mathbf{r}, t) = 0 \\ x_s(\mathbf{r}, t) = (1 + \sin(2\pi f_0 t)) \mathbb{1}_{\partial\Omega_s}(\mathbf{r}) \end{cases} \quad (\text{A III-9})$$

$$\Longleftrightarrow$$

$$\begin{cases} \widehat{X}_s(\mathbf{r}, f) = 0 \\ \widehat{x}_s(\mathbf{r}, f) = \left( \delta_f + \frac{1}{2i} (\delta_{f-f_0} - \delta_{f+f_0}) \right) \mathbb{1}_{\partial\Omega_s}(\mathbf{r}) \end{cases} \quad (\text{A III-10})$$

with  $f_0 = 100\text{MHz}$ , such that a period equal to 10 ns. Equation (A III-4) is solved for  $f = 0, f_0$  and  $-f_0$  with source terms from eq. (A III-10) and with the  $\mathcal{P}_1$  Lagrange finite element

method. Finally, the time-dependent solution at a given time is given by :

$$\varphi(\mathbf{r}, t) = \Re[\varphi(\mathbf{r}, 0) + \exp(2i\pi f_0 t) \widehat{\varphi}(\mathbf{r}, f_0) + \exp(-2i\pi f_0 t) \widehat{\varphi}(\mathbf{r}, -f_0)] \quad (\text{A III-11})$$

It is worth mentioning that for source terms such as eq. (A III-10), only stationary part of the photon density solutions and imaginary parts of the photon density solution for frequency  $-f_0$  contribute to the solution  $\varphi(\mathbf{r}, t)$ , eq. (A III-11). As a result, only two PDEs are solved for computing solution  $\varphi(\mathbf{r}, t)$  in order to increase the efficiency of the numerical scheme for the solution of the forward model. Figure III-1 depicts solutions obtained at twelve different time steps between 0 and 10ns with source terms equal to eqs. (A III-9)-(A III-10) and  $\partial\Omega_s(\mathbf{r}) = \{x \in [2/3; 4/3], y = 0\}$ . An inclusion with less absorbing and scattering properties, which was introduced at the south-west of the domain, can be guessed.

Let  $\varphi_a$  be the function

$$\varphi_a(\mathbf{r}, t) = \mathbf{r}_x^3 \cos(2\pi f_0 t), \quad \forall \mathbf{r} \in \Omega, \forall t > 0. \quad (\text{A III-12})$$

$\varphi_a$  is the analytical solution of eq. (A III-1) considering :

$$\begin{cases} X_s(\mathbf{r}, t) &= -\frac{2\pi f_0}{c} \mathbf{r}_x^3 \sin(2\pi f_0 t) - 6D(\mathbf{r}) \mathbf{r}_x \cos(2\pi f_0 t) + \kappa(\mathbf{r}) \varphi_a(\mathbf{r}, t) \\ x_s(\mathbf{r}, t) &= \gamma \varphi_a(\mathbf{r}, t) + \frac{3AD}{2} \mathbf{r}_x^2 \cos(2\pi f_0 t) \mathbf{n}_x \end{cases} \quad (\text{A III-13})$$

The frequency formulation is given by :

$$\begin{cases} \widehat{\varphi}_a(\mathbf{r}, f) &= \frac{1}{2} \mathbf{r}_x^3 (\delta_{f-f_0} + \delta_{f+f_0}) \\ \widehat{X}_s(\mathbf{r}, f) &= -\frac{\pi f_0}{c} \mathbf{r}_x^3 (\delta_{f-f_0} - \delta_{f+f_0}) - 3D(\mathbf{r}) \mathbf{r}_x (\delta_{f-f_0} + \delta_{f+f_0}) + \kappa(\mathbf{r}) \widehat{\varphi}_a(\mathbf{r}, f) \\ \widehat{x}_s(\mathbf{r}, f) &= \gamma \widehat{\varphi}_a(\mathbf{r}, f) + \frac{3AD}{2} \mathbf{r}_x^2 \mathbf{n}_x (\delta_{f-f_0} + \delta_{f+f_0}) \end{cases} \quad (\text{A III-14})$$

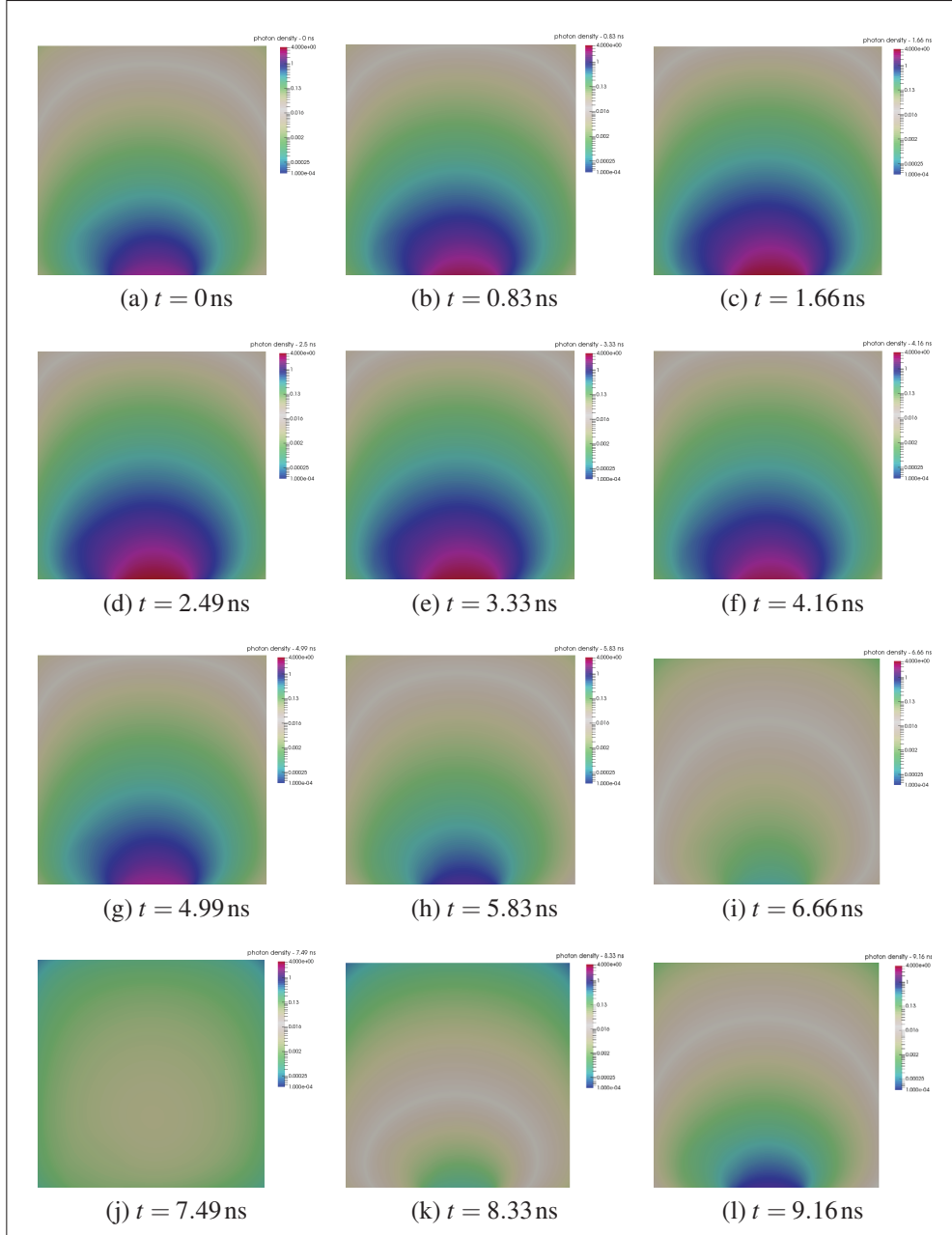


Figure-A III-1 Numerical solution of the photon density  $\varphi(\mathbf{r}, t)$  at different time steps ( $1/f = 10\text{ns}$ ).

Numerical solutions of eq. (A III-4) with such source terms and  $\kappa = 0.1\text{cm}^{-1}$ ,  $\sigma = 10\text{cm}^{-1}$ , are compared to the analytical solution, eq. (A III-12). Figure III-2 depicts, in logarithmic scales, the evolution of the  $L^2$  and  $H^1$  norms according to the dimension of the mesh size. Classical linear relations, in logarithmic scales, for the  $L^2$  and  $H^1$  norms and for  $\mathcal{P}_1$  Lagrange

finite elements are obtained. Following fits can be computed :

$$e_h^{L^2} \approx -2.4h^{2.004} \quad ; \quad e_h^{H^1} \approx 0.32h^{1.000} \quad (\text{A III-15})$$

where  $h$  and  $e_h$ , which represent  $x$  and  $y$  coordinates in the Figure III-2, are the characteristic length of the mesh and the error between the analytical and numerical solutions at time  $t = 1$  ns. Note that the theoretical study for such an equation and the use of  $\mathcal{P}_1$  Lagrange finite elements predicts a rate of convergence of 2 and 1 with respect to the norms  $L^2$  and  $H^1$ , respectively (Allaire et Craig, 2007). The present numerical study obtains a rate of convergence of 2.004 and 1.000 with respect to the norms  $L^2$  and  $H^1$ , respectively, which validates the numerical scheme for the solution of the diffuse approximation.

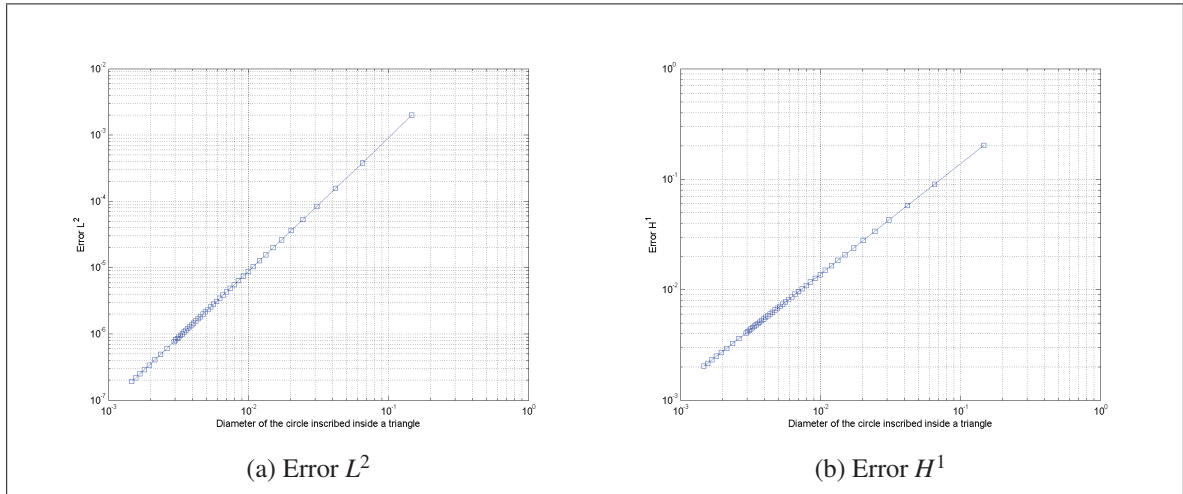


Figure-A III-2 Evolution of the  $L^2$  and  $H^1$  norms according to the dimension of the mesh size.





## ANNEXE IV

### LISTE DES PUBLICATIONS EN LIEN AVEC LA THÈSE

#### Articles acceptés dans des revues scientifiques avec comité de lecture

1. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. 2015. « Regularization opportunities for the diffuse optical tomography problem ». (98) 1–23. *International Journal of Thermal Sciences*. <http://doi.org/10.1016/j.ijthermalsci.2015.06.015>.
2. Dubot F., Y. Favennec, B. Rousseau, Y. Jarny, et D. R. Rousse. 2015. « Quasi-optimal Tikhonov penalization and parameterization coarseness in space-dependent function estimation ». *Inverse Problems in Science and Engineering*.  
<http://doi.org/10.1080/17415977.2015.1047362>.
3. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. 2015. « A wavelet multi-scale method for the inverse problem of diffuse optical tomography ». (289) 267–281. *Journal of Computational and Applied Mathematics*.  
<http://dx.doi.org/10.1016/j.cam.2015.01.023>.

#### Articles soumis dans des revues scientifiques avec comité de lecture

1. Dubot F., Y. Favennec, D. Le Hardy, B. Rousseau, et D. R. Rousse. 2015. « Space-dependent Sobolev gradients as a regularization for inverse radiative transfer problems ». *Journal of Quantitative Spectroscopy & Radiative Transfer*.

### Conférences internationales avec comité de lecture

1. Dubot F., O. Balima, Y. Favennec, et D. R. Rousse. April 2012. « The Sobolev gradient regularization strategy for optimal tomography coupled with a finite element formulation of the radiative transfer equation ». *6th conference on Inverse Problems, Control and Shape Optimization (PICO F'12)*. (École Polytechnique, Palaiseau, France 2012).
2. Dubot F., O. Balima, Y. Favennec, D. R. Rousse, et B. Rousseau. September 2012. « Filtering study of diffuse optical tomography reconstruction ». *Numerical Heat Transfer (NHT)*. (Wroclaw, Poland 2012).
3. Dubot F., O. Balima, Y. Favennec, B. Rousseau, et D. R. Rousse. June 2013. « Approximation diffuse de la propagation du rayonnement : un algorithme d'inversion appliqué à la détermination des propriétés radiatives ». *XIème Colloque Interuniversitaire Franco-Québécois sur la Thermique des Systèmes (CIFQ)*. (Reims, France 2013).
4. Favennec Y., F. Dubot, B. Rousseau, et D. R. Rousse. June 2013. « Mixing regularization tools for enhancing regularity in optical tomography applications ». *4th Inverse problems, design and optimization symposium (IPDO)*. (Albi, France 2013).
5. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. June 2014. « Wavelet-based parameterization of control variables in diffuse optical tomography ». *6th International Conference on Advanced Computational Methods in ENgineering (ACOMEN)*. (Ghent, Belgium 2014).
6. Rousseau B., Y. Favennec, S. Guevelou, F. Dubot, et D. R. Rousse. July 2014. « Infrared tomography : towards a novel methodology to investigate the volumetric radiative properties of heterogeneous materials ». *12th International Conference on Quantitative InfraRed Thermography (QIRT)*. (Bordeaux, France 2014).
7. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. April 2015. « Determination of space-dependent radiative properties in diffuse optical tomography using a wavelet multi-scale method ». *Computational Thermal Radiation in Participating Media V (CTRPM)*. (Albi, France 2015).

8. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. June 2015. « Une méthode inverse multi-échelle basée sur les ondelettes pour la résolution du problème de tomographie optique diffuse ». *XIIème Colloque Interuniversitaire Franco-Québécois sur la Thermique des Systèmes (CIFQ)*. (Sherbrooke, Québec, Canada 2015).

### **Conférences nationales avec comité de lecture**

1. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. May 2013. « Optimisation de l'algorithme du L-BFGS en vue de la reconstruction simultanée de deux fonctions dans le cadre de la tomographie optique ». *6e Biennale Française des Mathématiques Appliquées et Industrielles (Congrès SMAI 2013)*. (Seignosses, France 2013).
2. Dubot F., Y. Favennec, B. Rousseau, et D. R. Rousse. June 2014. « Paramétrisation des variables de contrôle et méthodes de recherche linéaire dans un code d'inversion de l'approximation de diffusion basé sur le L-BFGS ». *Congrès de la Société Française de Thermique (SFT)*. (Lyon, France 2014).



## BIBLIOGRAPHIE

- Addam, M. 2009. « Approximation du problème de diffusion en tomographie optique et problème inverse ». PhD thesis, Université du Littoral Côte d'Opale. <<https://tel.archives-ouvertes.fr/tel-00579257>>.
- Addison, P. S., 2010. *The illustrated wavelet transform handbook : introductory theory and applications in science, engineering, medicine and finance*.
- Alifanov, O. 1974. « Solution of an inverse problem of heat conduction by iteration methods ». *Journal of Engineering Physics and Thermophysics*, vol. 26, n° 4, p. 471–476.
- Allaire, G. et A. Craig, 2007. *Numerical analysis and optimization*.
- Alonzo, C. A., M. C. Gather, J. W. Kang, G. Scarcelli, et Y. Seok-Hyun, 2011. *Handbook of biomedical optics*, chapter 4 : Light sources, detectors, and irradiation guidelines, p. 49–65. CRC Press.
- Amestoy, P. R., A. Guermouche, J.-Y. L'Excellent, et S. Pralet. 2006. « Hybrid scheduling for the parallel solution of linear systems ». *Parallel computing*, vol. 32, n° 2, p. 136–156.
- Antoniou, A. et W.-S. Lu, 2007. *Practical optimization : algorithms and engineering applications*.
- Arridge, S. R., M. Schweiger, et J. C. Schotland, 2011. *Handbook of biomedical optics*, chapter 17 : Inverse Models of Light Transport, p. 319–336. CRC Press.
- Arridge, S. et M. Schweiger. 1998. « A gradient-based optimisation scheme for optical tomography ». *Optics Express*, vol. 2, n° 6, p. 213–226.
- Arridge, S. R. 1999. « Optical tomography in medical imaging ». *Inverse problems*, vol. 15, n° 2, p. R41.
- Arridge, S. R. et W. R. Lionheart. 1998. « Nonuniqueness in diffusion-based optical tomography ». *Optics Letters*, vol. 23, n° 11, p. 882–884.
- Arridge, S. R. et M. Schweiger. 1995. « Photon-measurement density functions. Part 2 : Finite-element-method calculations ». *Applied Optics*, vol. 34, n° 34, p. 8026–8037.
- Arridge, S. R., H. Dehghani, M. Schweiger, et E. Okada. 2000. « The finite element model for the propagation of light in scattering media : a direct method for domains with nonscattering regions ». *Medical physics*, vol. 27, n° 1, p. 252–264.
- Arridge, S. R., O. Dorn, V. Kolehmainen, M. Schweiger, et A. Zacharopoulos. 2008. « Parameter and structure reconstruction in optical tomography ». In *Journal of Physics : Conference Series*. p. 012001. IOP Publishing.
- Arridge, S., J. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, et M. Vauhkonen. 2006. « Approximation errors and model reduction with an application in optical diffusion tomography ». *Inverse Problems*, vol. 22, n° 1, p. 175.
- Asllanaj, F., A. Addoum, et S. Contassot-Vivier. 2015. « Detection of tumor-like inclusions embedded within human liver tissue using a short-pulsed near-infrared laser beam : Parallel simulations with radiative transfer equation ». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 165, p. 1 - 11.

- Ayachit, U., 2015. *The ParaView Guide : A Parallel Visualization Application*.
- Azizi, L., K. Zarychta, D. Etti, E. Tinet, et J.-M. Tualle. 2009. «Ultimate spatial resolution with diffuse optical tomography». *Optics express*, vol. 17, n° 14, p. 12132–12144.
- Baillis, D. et J.-F. Sacadura. 2000. «Thermal radiation properties of dispersed media : theoretical prediction and experimental characterization». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 67, n° 5, p. 327–363.
- Balima, O., Y. Favennec, F. Dubot, et D. Rousse. 2012. «Finite elements parameterization of optical tomography with the radiative transfer equation in frequency domain». In *Journal of Physics : Conference Series*. p. 012022. IOP Publishing.
- Balima, O., Y. Favennec, et D. Rousse. 2013. «Optical tomography reconstruction algorithm with the finite element method : An optimal approach with regularization tools». *Journal of Computational Physics*, vol. 251, p. 461–479.
- Balima, O., J. Boulanger, A. Charette, et D. Marceau. 2011a. «New developments in frequency domain optical tomography. Part II : Application with a L-BFGS associated to an inexact line search». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 112, n° 7, p. 1235–1240.
- Balima, O., J. Boulanger, A. Charette, et D. Marceau. 2011b. «New developments in frequency domain optical tomography. Part I : Forward model and gradient computation». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 112, n° 7, p. 1229–1234.
- Barbour, R. L., H. L. Graber, et Y. Pei. 2003. «Normalized-constraint algorithm for minimizing inter-parameter crosstalk in imaging of scattering media». US Patent 6,662,128.
- Bashkatov, A. N., E. A. Genina, et V. V. Tuchin, 2011. *Handbook of biomedical optics*, chapter 5 : Tissue optical properties, p. 67–100. CRC Press.
- Bassani, M., F. Martelli, G. Zaccanti, et D. Contini. 1997. «Independence of the diffusion coefficient from absorption : experimental and numerical evidence». *Optics letters*, vol. 22, n° 12, p. 853–855.
- Beilina, L. et M. V. Klibanov. 2008. «A globally convergent numerical method for a coefficient inverse problem». *SIAM Journal on Scientific Computing*, vol. 31, n° 1, p. 478–509.
- Beilina, L. et M. V. Klibanov, 2012. *Approximate global convergence and adaptivity for coefficient inverse problems*.
- Björck, A., 1996. *Numerical methods for least squares problems*.
- Boas, D. A., A. M. Dale, et M. A. Franceschini. 2004. «Diffuse optical imaging of brain activation : approaches to optimizing image sensitivity, resolution, and accuracy». *Neuroimage*, vol. 23, p. S275–S288.
- Boffety, M. 2010. «Étude quantitative de la tomographie optique diffuse de luminescence. Application à la localisation de sources en imagerie moléculaire». PhD thesis, École Centrale Paris.
- Boulanger, J., O. Balima, et A. Charette. 2011. «Refractive index imaging from radiative transfer equation-based reconstruction algorithm : Fundamentals». *J. Quant. Spectrosc. Radiat. Transfer*, vol. 112, p. 1222–1228.

- Bourquin, F. et A. Nassiopoulou. 2011. «Inverse reconstruction of initial and boundary conditions of a heat transfer problem with accurate final state». *International Journal of Heat and Mass Transfer*, vol. 54, p. 3749-3760.
- Brattka, V. et A. Yoshikawa. 2006. «Towards computability of elliptic boundary value problems in variational formulation». *Journal of Complexity*, vol. 22, n° 6, p. 858–880.
- Cao, N., A. Nehorai, et M. Jacobs. 2007. «Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm». *Optics express*, vol. 15, n° 21, p. 13695–13708.
- Céa, J., 1971. *Optimisation, théorie et algorithmes*.
- Chai, J. C., H. S. Lee, et S. V. Patankar. 1994. «Finite volume method for radiation heat transfer». *Journal of thermophysics and heat transfer*, vol. 8, n° 3, p. 419–425.
- Chandrasekhar, S., 1960. *Radiative transfer*.
- Charette, A., J. Boulanger, et H. K. Kim. 2008. «An overview on recent radiation transport algorithm development for optical tomography imaging». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 109, n° 17–18, p. 2743–2766.
- Charette, A., O. Balima, et J. Boulanger. June 2011. «Le Rayonnement en Milieux Diffusants : Application à la Tomographie Optique». In *Colloque Interuniversitaire Franco-Québécois*. (Chicoutimi, Canada 2011).
- Chavent, G., 2010. *Nonlinear least squares for inverse problems : theoretical foundations and step-by-step guide for applications*.
- Chen, N. G. et J. Bai. 1998. «Monte Carlo approach to modeling of boundary conditions for the diffusion equation». *Physical review letters*, vol. 80, n° 24, p. 5321.
- Cheong, W., S. Prahl, et A. Welch. 1990. «A Review of the Optical Properties of the Biological Tissues». *IEEE Journal of Quantum Electronics*, vol. 26, p. 2166–2185.
- Christiansen, C. 1883. «Absolute Bestimmung des Emissions-und Absorptionsvermögens für Wärme». *Annalen der Physik*, vol. 255, n° 6, p. 267–283.
- Chu, C., S. Churchill, et G. Clark, 1957. *Tables of angular distribution coefficients for light-scattering by spheres*.
- Colton, D. L. et R. Kress, 1992. *Inverse acoustic and electromagnetic scattering theory*, volume 93.
- Comité consultatif de la Société canadienne du cancer. Mai 2014. «Statistiques canadiennes sur le cancer 2014». Toronto (Ontario) : Société canadienne du cancer. <<http://www.cancer.ca/~media/cancer.ca/CW/publications/Canadian%20Cancer%20Statistics%202014/Canadian-Cancer-Statistics-2014-FR.pdf>>.
- Danaila, I. et P. Kazemi. 2010. «A new Sobolev gradient method for direct minimization of the Gross-Pitaevskii energy with rotation». *Siam Journal of scientific computing*, vol. 32, n° 5, p. 2447-2467.
- Dehghani, H., S. Srinivasan, B. W. Pogue, et A. Gibson. 2009a. «Numerical modelling and image reconstruction in diffuse optical tomography». *Philosophical Transactions of the*



- Royal Society A : Mathematical, Physical and Engineering Sciences*, vol. 367, n° 1900, p. 3073–3093.
- Dehghani, H., S. R. Arridge, M. Schweiger, et D. T. Delpy. 2000. «Optical tomography in the presence of void regions». *JOSA A*, vol. 17, n° 9, p. 1659–1670.
- Dehghani, H., M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, et K. D. Paulsen. 2009b. «Near infrared optical tomography using NIR-FAST : Algorithm for numerical model and image reconstruction». *Communications in numerical methods in engineering*, vol. 25, n° 6, p. 711–732.
- Dehghani, H., B. R. White, B. W. Zeff, A. Tizzard, et J. P. Culver. 2009c. «Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography». *Applied optics*, vol. 48, n° 10, p. D137–D143.
- Dennis, J. et R. Schnabel, 1996. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*.
- Ding, L., B. Han, et J.-q. Liu. 2009. «A wavelet multiscale method for inversion of Maxwell equations». *Applied Mathematics and Mechanics*, vol. 30, p. 1035–1044.
- Dubot, F., Y. Favennec, B. Rousseau, Y. Jarny, et D. Rousse. 2015a. «Quasi-optimal Tikhonov penalization and parameterization coarseness in space-dependent function estimation». *Inverse Problems in Science and Engineering*.
- Dubot, F., Y. Favennec, B. Rousseau, et D. R. Rousse. June 2014. «Paramétrisation des variables de contrôle et méthodes de recherche linéaire dans un code d'inversion de l'approximation de diffusion basé sur le L-BFGS». In *Comptes-Rendus Congrès de la Société Française de Thermique*. (Lyon, France 2014).
- Dubot, F., Y. Favennec, B. Rousseau, et D. R. Rousse. 2015b. «Regularization opportunities for the diffuse optical tomography problem». *International Journal of Thermal Sciences*, vol. 98, p. 1–23.
- Dubot, F., Y. Favennec, B. Rousseau, et D. R. Rousse. 2015c. «A wavelet multi-scale method for the inverse problem of diffuse optical tomography». *Journal of Computational and Applied Mathematics*, vol. 289, p. 267–281.
- Dulikravich, G. et T. Martin. 1996. «Inverse shape and boundary condition problems and optimization in heat conduction». *Advances in Numerical Heat Transfer*, vol. 1, p. 381–426.
- Eason, G., A. Veitch, R. Nisbet, et F. Turnbull. 1978. «The theory of the back-scattering of light by blood». *Journal of Physics D : Applied Physics*, vol. 11, n° 10, p. 1463.
- Elaloufi, R., R. Carminati, et J.-J. Greffet. 2002. «Time-dependent transport through scattering media : from radiative transfer to diffusion». *Journal of Optics A : Pure and Applied Optics*, vol. 4, n° 5, p. S103.
- Favennec, Y., P. Le Masson, et Y. Jarny. June 2011. Lecture 7 : Optimization methods for non linear estimation or function estimation. Eurotherm Spring School METTI 2011 : Thermal measurements and inverse techniques, Roscoff. <<http://www.sft.asso.fr/document.php?pagendx=12299&project=sft>>.



- Favennec, Y., D. Le Hardy, F. Dubot, B. Rousseau, et D. R. Rousse. April 1–3 2015. «Some speed-up strategies for solving inverse radiative transfer problems». In *Eurotherm 105 – Computational Thermal Radiation in Participating Media V.* (Albi 2015).
- Favennec, Y., F. Dubot, B. Rousseau, et D. R. Rousse. June 26–28 2013. «Mixing regularization tools for enhancing regularity in optical tomography applications». In *IPDO 2013 : 4th Inverse problems, design and optimization symposium.* (Albi 2013).
- Firbank, M., S. R. Arridge, M. Schweiger, et D. T. Delpy. 1996. «An investigation of light transport through scattering bodies with non-scattering regions». *Physics in medicine and biology*, vol. 41, n° 4, p. 767.
- Fiveland, W. 1984. «Discrete-ordinates solutions of the radiative transport equation for rectangular enclosures». *Journal of Heat Transfer*, vol. 106, n° 4, p. 699–706.
- Fiveland, W. 1988. «Three-dimensional radiative heat-transfer solutions by the discrete-ordinates method». *Journal of Thermophysics and Heat Transfer*, vol. 2, n° 4, p. 309–316.
- Flannery, B. P., W. H. Press, S. A. Teukolsky, et W. Vetterling. 1992. «Numerical recipes in C». *Press Syndicate of the University of Cambridge, New York*, vol. 24.
- Fletcher, R. 1980. «Practical methods of optimization, Vol. 1, Unconstrained optimization». *British Library Cataloguing in Publication Data.*
- Francoeur, M., R. Vaillon, et D. R. Rousse. 2005. «Theoretical analysis of frequency and time-domain methods for optical characterization of absorbing and scattering media». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 93, n° 1, p. 139–150.
- Fu, H., B. Han, et H. Liu. 2013. «A wavelet multiscale iterative regularization method for the parameter estimation problems of partial differential equations». *Neurocomputing*, vol. 104, p. 138–145.
- Fu, H., B. Han, et G. Gai. 2007. «A wavelet multiscale–homotopy method for the inverse problem of two-dimensional acoustic wave equation». *Applied mathematics and computation*, vol. 190, n° 1, p. 576–582.
- Gao, H. et H. Zhao. 2009. «A Fast-Forward Solver of Radiative Transfer Equation». *Transport Theory and Statistical Physics*, vol. 38, p. 149–192.
- Geuzaine, C. et J.-F. Remacle. 2009. «Gmsh : A 3-D finite element mesh generator with built-in pre-and post-processing facilities». *International Journal for Numerical Methods in Engineering*, vol. 79, n° 11, p. 1309–1331.
- Gibson, A., J. Hebden, et S. R. Arridge. 2005. «Recent advances in diffuse optical imaging». *Physics in medicine and biology*, vol. 50, n° 4, p. R1.
- Golub, G. H., M. Heath, et G. Wahba. 1979. «Generalized cross-validation as a method for choosing a good ridge parameter». *Technometrics*, vol. 21, n° 2, p. 215–223.
- Goncharskii, A., A. S. Leonov, et A. G. Yagola. 1973. «A generalized discrepancy principle». *USSR Computational Mathematics and Mathematical Physics*, vol. 13, n° 2, p. 25–37.
- Groenhuis, R., H. A. Ferwerda, et J. Ten Bosch. 1983. «Scattering and absorption of turbid materials determined from reflection measurements. 1 : Theory». *Applied Optics*, vol. 22, n° 16, p. 2456–2462.

- Gu, X., Y. Xu, et H. Jiang. 2003. « Mesh-based enhancement schemes in diffuse optical tomography ». *Medical physics*, vol. 30, n° 5, p. 861–869.
- Hadamard, J. 1902. « Sur les problèmes aux dérivées partielles et leur signification physique ». *Princeton university bulletin*, vol. 13, n° 49–52, p. 28.
- Hajihashemi, M. R., S. R. Grobmyer, S. Z. Al-Quran, et H. Jiang. 2012. « Noninvasive evaluation of nuclear morphometry in breast lesions using multispectral diffuse optical tomography ». *PloS one*, vol. 7, n° 9, p. e45714.
- Hammersley, J. M. et D. C. Handscomb, 1964. *Monte carlo methods*, volume 1.
- Hansen, P. C. et D. P. O’Leary. 1993. « The use of the L-curve in the regularization of discrete ill-posed problems ». *SIAM Journal on Scientific Computing*, vol. 14, n° 6, p. 1487–1503.
- Harrach, B. 2009. « On uniqueness in diffuse optical tomography ». *Inverse problems*, vol. 25, n° 5, p. 055010.
- Haskell, R. C., L. O. Svaasand, T.-T. Tsay, T.-C. Feng, M. S. McAdams, et B. J. Tromberg. 1994. « Boundary conditions for the diffusion equation in radiative transfer ». *J. Opt. Soc. Am. A*, vol. 11, n° 10, p. 2727–2741.
- Hayashi, T., Y. Kashio, et E. Okada. 2003. « Hybrid Monte Carlo-diffusion method for light propagation in tissue with a low-scattering region ». *Applied optics*, vol. 42, n° 16, p. 2888–2896.
- He, Y. et B. Han. 2009. « A wavelet adaptive-homotopy method for inverse problem in the fluid-saturated porous media ». *Applied Mathematics and Computation*, vol. 208, n° 1, p. 189–196.
- Hecht, F. 2012. « New development in FreeFem++ ». *Journal of Numerical Mathematics*, vol. 20, n° 3-4, p. 251–265.
- Hecht, F., S. Auliac, O. Pironneau, J. Morice, A. Le Hyaric, et K. Ohtsuka. 2014. « FreeFem++ manual ». <<http://www.freefem.org/ff++/ftp/freefem++doc.pdf>>.
- Heino, J. et E. Somersalo. 2002. « Estimation of optical absorption in anisotropic background ». *Inverse Problems*, vol. 18, n° 3, p. 559.
- Hensel, E. et R. Hills. 1989. « Steady-state two-dimensional inverse heat conduction ». *Numerical Heat Transfer, Part B Fundamentals*, vol. 15, n° 2, p. 227–240.
- Heney, L. et J. Greenstein. 1941. « Diffuse Radiation in the Galaxy ». *Astrophysical Journal*, vol. 93, p. 70–83.
- Hielscher, A. H., H. K. Kim, et A. D. Klose, 2011. *Handbook of biomedical optics*, chapter 16 : Forward models of light transport in biological tissue, p. 319–336. CRC Press.
- Hielscher, A., A. Bluestone, G. Abdoulaev, A. Klose, J. Lasker, M. Stewart, U. Netz, et J. Beuthan. 2002. « Near-infrared diffuse optical tomography ». *Disease markers*, vol. 18, n° 5-6, p. 313–337.
- Hielscher, A. H. 2005. « Optical tomographic imaging of small animals ». *Current opinion in biotechnology*, vol. 16, n° 1, p. 79–88.

- Hielscher, A. H. et R. E. Alcouffe. 1996. «Nondiffusive photon migration in homogeneous and heterogeneous tissues». In *BiOS Europe'96*. p. 22–30. International Society for Optics and Photonics.
- Hielscher, A. H., S. L. Jacques, L. Wang, et F. Tittel. 1995. «The influence of boundary conditions on the accuracy of diffusion theory in time-resolved reflectance spectroscopy of biological tissues». *Physics in Medicine and Biology*, vol. 40, n° 11, p. 1957.
- Hielscher, A. H., R. E. Alcouffe, et R. L. Barbour. 1998. «Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues». *Physics in Medicine and Biology*, vol. 43, n° 5, p. 1285.
- Hielscher, A. H., A. D. Klose, A. K. Scheel, B. Moa-Anderson, M. Backhaus, U. Netz, et J. Beuthan. 2004. «Sagittal laser optical tomography for imaging of rheumatoid finger joints». *Physics in Medicine and Biology*, vol. 49, n° 7, p. 1147.
- Hillman, E. 2002. «Experimental and theoretical investigations of near infrared tomographic imaging methods and clinical applications». PhD thesis, University College London.
- Howell, J. R. 1968. «Application of Monte Carlo to heat transfer problems». *Advances in heat transfer*, vol. 5, n° 1, p. 1–54.
- Howell, J. 1998. «The Monte Carlo method in radiative heat transfer». *Journal of Heat Transfer*, vol. 120, n° 3, p. 547–560.
- Howell, J., R. Siegel, et M. Mengüç, 2011. *Thermal Radiation Heat Transfer*. éd. 5.
- Jackson, J. et R. Fox. 1999. «Classical electrodynamics». *American Journal of Physics*, vol. 67, p. 841.
- Jagannath, R. P. K., P. K. Yalavarthy, et al. 2013. «Nonquadratic penalization improves near-infrared diffuse optical tomography». *J. Opt. Soc. Am. A*, vol. 30, n° 8, p. 1516–1523.
- Jarny, Y. June 2011. Lecture 9 : Inverse problems & regularized solutions. Eurotherm Spring School METTI 2011 : Thermal measurements and inverse techniques, Roscoff. <<http://www.sft.asso.fr/document.php?pagendx=12299&project=sft>>.
- Jeans, J. 1917. «The equations of radiative transfer of energy». *Monthly Notices of the Royal Astronomical Society*, vol. 78, n° 1, p. 28–36.
- Jiang, H., 2010. *Diffuse optical tomography : principles and applications*.
- Jin, G., Q. Liu, et X. Lv. 2015. «Inversion Study of Vertical Eddy Viscosity Coefficient Based on an Internal Tidal Model with the Adjoint Method». *Mathematical Problems in Engineering*, vol. 2015.
- Kaipio, J. et E. Somersalo, 2005. *Statistical and computational inverse problems*, volume 160.
- Keijzer, M., W. M. Star, et P. R. Storch. 1988. «Optical diffusion in layered media». *Applied Optics*, vol. 27, n° 9, p. 1820–1824.
- Khairi, M., S. Ibrahim, M. Yunus, et M. Faramarzi. 2012. «A Review on Applications of Optical Tomography in Industrial Process». *International Journal on Smart Sensing and Intelligent Systems*, vol. 5, n° 4, p. 767–798.

- Kim, H. K. et A. H. Hielscher. 2009. « A PDE-constrained SQP algorithm for optical tomography based on the frequency-domain equation of radiative transfer ». *Inverse Problems*, vol. 25, n° 1, p. 015010.
- Kim, H. K. et A. H. Hielscher. 2010. « A diffusion–transport hybrid method for accelerating optical tomography ». *Journal of Innovative Optical Health Sciences*, vol. 3, n° 04, p. 293–305.
- Kim, K. W., S. W. Baek, M. Y. Kim, et H. S. Ryou. 2004. « Estimation of emissivities in a two-dimensional irregular geometry by inverse radiation analysis using hybrid genetic algorithm ». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 87, n° 1, p. 1–14.
- Klose, A. October 2001. « Optical tomography based on the Equation of Radiative Transfer ». PhD thesis, Department of Physics Freie Universität, Berlin Germany.
- Klose, A. D. et A. H. Hielscher. 2003. « Quasi-Newton methods in optical tomographic image reconstruction ». *Inverse problems*, vol. 19, n° 2, p. 387.
- Klose, A. D., U. Netz, J. Beuthan, et A. H. Hielscher. 2002. « Optical tomography using the time-independent equation of radiative transfer—Part 1 : forward model ». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 72, n° 5, p. 691–713.
- Leff, D. R., O. J. Warren, L. C. Enfield, A. Gibson, T. Athanasiou, D. K. Patten, J. Hebden, G. Z. Yang, et A. Darzi. 2008. « Diffuse optical imaging of the healthy and diseased breast : a systematic review ». *Breast cancer research and treatment*, vol. 108, n° 1, p. 9–22.
- Lei, J., S. Liu, Z. Li, M. Sun, et X. Wang. 2011. « A multi-scale image reconstruction algorithm for electrical capacitance tomography ». *Applied Mathematical Modelling*, vol. 35, n° 6, p. 2585–2606.
- Li, B.-W., W.-Q. Tao, et R.-X. Liu. 1997. « Ray effect in ray tracing method for radiative heat transfer ». *International journal of heat and mass transfer*, vol. 40, n° 14, p. 3419–3426.
- Lions, J.-L., 1968. *Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles*. Paris : Dunod.
- Lions, J.-L. et P. Faurre, 1982. *Cours d'analyse numérique*.
- Liu, C.-S. 2014. « Optimal Algorithms and the BFGS Updating Techniques for Solving Unconstrained Nonlinear Minimization Problems ». *Journal of Applied Mathematics*, vol. 2014.
- Liu, D. et J. Nocedal. 1989. « On the limited memory BFGS method for large scale optimization ». *Mathematical Programming : Series A and B*, vol. 45, n° 3, p. 503–528.
- Liu, J. 1993. « A multiresolution method for distributed parameter estimation ». *SIAM Journal on Scientific Computing*, vol. 14, n° 2, p. 389–405.
- Mahan, J., 2002. *Radiation Heat Transfer : A Statistical Approach*.
- Mahaviev, W. 1997. « A numerical method utilizing weighted Sobolev descent to solve singular differential equations ». *Nonlinear World*, vol. 4, p. 435-455.

- Maillet, D., Y. Favennec, et P. Le Masson. June 2011. *Eurotherm spring school METTI 2011 : Thermal Measurements and Inverse Techniques*. Technical report. Roscoff : <http://www.sft.asso.fr>.
- Majid, A. et S. Sial. 2010. « Application of Sobolev gradient method to Poisson–Boltzmann system ». *Journal of Computational Physics*, vol. 229, p. 5742–5754.
- Mallat, S., 1999. *A wavelet tour of signal processing*.
- McBride, T., B. W. Pogue, U. L. Osterberg, et K. D. Paulsen. 2000. « Separation of absorption and scattering heterogeneities in NIR tomographic imaging of tissue ». In *Biomedical Optical Spectroscopy and Diagnostics*. p. TuC2. Optical Society of America.
- McBride, T. O., B. W. Pogue, S. Jiang, U. L. Oesterberg, et K. D. Paulsen. 2001a. « A parallel-detection frequency-domain near-infrared tomography system for hemoglobin imaging of the breast in vivo ». *Review of Scientific Instruments*, vol. 72, n° 3, p. 1817–1824.
- McBride, T. O., B. W. Pogue, S. Jiang, U. L. Österberg, K. D. Paulsen, et S. P. Poplack. 2001b. « Initial studies of in vivo absorbing and scattering heterogeneity in near-infrared tomographic breast imaging ». *Optics letters*, vol. 26, n° 11, p. 822–824.
- Modest, M. F., 2013. *Radiative heat transfer*.
- Mohebbi, F. et M. Sellier. 2014. « Aerodynamic Optimal Shape Design Based on Body-Fitted Grid Generation ». *Mathematical Problems in Engineering*, vol. 2014.
- Morozov, V. A., Z. Nashed, et A. Aries, 1984. *Methods for solving incorrectly posed problems*.
- Moura, L. 1998. « Identification des propriétés radiatives des matériaux semi-transparents diffusants en situation de non-symétrie azimutale du champ radiatif ». PhD thesis, Institut National des Sciences Appliquées de Lyon.
- Murio, D. A., 1993. *The mollification method and the numerical solution of ill-posed problems*.
- Nakai, T., G. Nishimura, K. Yamamoto, et M. Tamura. 1997. « Expression of optical diffusion coefficient in high-absorption turbid media ». *Physics in medicine and biology*, vol. 42, n° 12, p. 2541.
- Nassiopoulos, A. et F. Bourquin. 2010. « Fast three-dimensional temperature reconstruction ». *Computer Methods in Applied Mechanics and Engineering*, vol. 199, p. 3169 - 3178.
- Nath, S. K., R. Vasu, et M. Pandit. 1999. « Wavelet based compression and denoising of optical tomography data ». *Optics communications*, vol. 167, n° 1, p. 37–46.
- Niu, H., P. Guo, L. Ji, Q. Zhao, et T. Jiang. 2008. « Improving image quality of diffuse optical tomography with a projection-error-based adaptive regularization method ». *Optics express*, vol. 16, n° 17, p. 12423–12434.
- Nocedal, J. 1980. « Updating quasi-Newton matrices with limited storage ». *Mathematics of computation*, vol. 35, n° 151, p. 773–782.
- O’Leary, D. P. 2001. « Near-optimal parameters for Tikhonov and other regularization methods ». *SIAM Journal on scientific computing*, vol. 23, n° 4, p. 1161–1171.
- O’Sullivan, T. D., A. E. Cerussi, D. J. Cuccia, et B. J. Tromberg. 2012. « Diffuse optical imaging using spatially and temporally modulated light ». *Journal of biomedical optics*, vol. 17, n° 7, p. 0713111–07131114.



- Parwani, A. K., P. Talukdar, et P. Subbarao. 2013. «Performance evaluation of hybrid differential evolution approach for estimation of the strength of a heat source in a radiatively participating medium». *International Journal of Heat and Mass Transfer*, vol. 56, n° 1, p. 552–560.
- Patterson, M. S., B. Chance, et B. C. Wilson. 1989. «Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties». *Applied optics*, vol. 28, n° 12, p. 2331–2336.
- Paulsen, K. D. et H. Jiang. 1995. «Spatially varying optical property reconstruction using a finite element diffusion equation approximation». *Medical Physics*, vol. 22, n° 6, p. 691–701.
- Pei, Y., H. Graber, et R. Barbour. 2001. «Normalized-constraint algorithm for minimizing inter-parameter crosstalk in DC optical tomography». *Optics express*, vol. 9, n° 2, p. 97–109.
- Pierrat, R. 2007. «Propagation et émission du rayonnement en milieu diffusant. Application à l'imagerie des milieux complexes». PhD thesis, Ecole Centrale Paris.
- Piron, V. et J.-P. L'Huillier. 2006. «Detection of heterogeneities embedded within a turbid slab media using time-and frequency-domain methods : application to the mammography». *Lasers in medical science*, vol. 21, n° 2, p. 67–73.
- Pogue, B., M. Testorf, T. McBride, U. Osterberg, et K. Paulsen. 1997. «Instrumentation and design of a frequency-domain diffuse optical tomography imager for breast cancer detection». *Optics express*, vol. 1, n° 13, p. 391–403.
- Protas, B. 2008. «Adjoint-based optimization of PDE systems with alternative gradients». *Journal of Computational Physics*, vol. 227, p. 6490–6510.
- Protas, B., T. Bewley, et G. Hagen. 2004. «A computational framework for the regularization of adjoint analysis in multiscale PDE systems». *Journal of Computational Physics*, vol. 195, n° 1, p. 49–89.
- Raithby, G. et E. Chui. 1990. «A finite-volume method for predicting a radiant heat transfer in enclosures with participating media». *Journal of Heat Transfer*, vol. 112, n° 2, p. 415–423.
- Razzaque, M., D. Klein, et J. Howell. 1983. «Finite element solution of radiative heat transfer in a two-dimensional rectangular enclosure with gray participating media». *Journal of heat transfer*, vol. 105, n° 4, p. 933–936.
- Razzaque, M., J. Howell, et D. Klein. 1984. «Coupled radiative and conductive heat transfer in a two-dimensional rectangular enclosure with gray participating media using finite elements». *Journal of heat transfer*, vol. 106, n° 3, p. 613–619.
- Reginska, T. 1996. «A regularization parameter in discrete ill-posed problems». *SIAM Journal on Scientific Computing*, vol. 17, n° 3, p. 740–749.
- Ren, K., G. Bal, et A. H. Hielscher. 2006. «Frequency domain optical tomography based on the equation of radiative transfer». *SIAM Journal on Scientific Computing*, vol. 28, n° 4, p. 1463–1489.

- Renka, R. 2009. «Image segmentation with a Sobolev gradient method». *Nonlinear Analysis*, vol. 71, p. 774–780.
- Riley, J., H. Dehghani, M. Schweiger, S. Arridge, J. Ripoll, et M. Nieto-Vesperinas. 2000. «3D optical tomography in the presence of void regions». *Optics express*, vol. 7, n° 13, p. 462–467.
- Rousse, D. R. 1994. «Numerical Predictions of Multidimensional Conduction, Convection and Radiation Heat Transfer in Participating Media». PhD thesis, McGill University.
- Rousse, D. R. 2000. «Numerical predictions of two-dimensional conduction, convection, and radiation heat transfer. I. Formulation». *International journal of thermal sciences*, vol. 39, n° 3, p. 315–331.
- Rousse, D. R., G. Gautier, et J.-F. Sacadura. 2000. «Numerical predictions of two-dimensional conduction, convection, and radiation heat transfer. II. Validation». *International journal of thermal sciences*, vol. 39, n° 3, p. 332–353.
- Rousseau, B., J. Rolland, P. Echegut, E. Brun, et J. Vicente. 2012. «Numerical prediction of the radiative behavior of metallic foams from the microscopic to macroscopic scale». In *Journal of Physics : Conference Series*. p. 012003. IOP Publishing.
- Rousseau, B., D. de Sousa Meneses, P. Echegut, M. Di Michiel, et J.-F. Thovert. 2007. «Prediction of the thermal radiative properties of an X-ray  $\mu$ -tomographed porous silica glass». *Applied optics*, vol. 46, n° 20, p. 4266–4276.
- Rousseau, B., S. Guevelou, G. Domingues, J. Vicente, C. Caliot, et G. Flamant. 2013. «Prediction of the radiative properties of reconstructed alpha-SiC foams used for concentrated solar applications». *MRS Online Proceedings Library*, vol. 1545, p. mrss13–1545.
- Sacadura, J.-F. 2011. «Thermal radiative properties of complex media : theoretical prediction versus experimental identification». *Heat Transfer Engineering*, vol. 32, n° 9, p. 754–770.
- Schweiger, M., S. Arridge, M. Hiraoka, et D. Delpy. 1995. «The finite element method for the propagation of light in scattering media : boundary and source conditions». *Medical physics*, vol. 22, n° 11, p. 1779–1792.
- Schweiger, M. et S. R. Arridge. 1999. «Application of temporal filters to time resolved data in optical tomography». *Physics in medicine and biology*, vol. 44, n° 7, p. 1699.
- Schweiger, M., S. R. Arridge, et I. Nissilä. 2005. «Gauss–Newton method for image reconstruction in diffuse optical tomography». *Physics in medicine and biology*, vol. 50, n° 10, p. 2365.
- Shan, H., M. V. Klibanov, H. Liu, N. Pantong, et J. Su. 2008a. «Numerical implementation of the convexification algorithm for an optical diffusion tomograph». *Inverse Problems*, vol. 24, n° 2, p. 025006.
- Shan, H., M. V. Klibanov, J. Su, N. Pantong, et H. Liu. 2008b. «A globally accelerated numerical method for optical tomography with continuous wave source». *Journal of Inverse and Ill-posed Problems*, vol. 16, n° 8, p. 763–790.

- Shaw, C. B. et P. K. Yalavarthy. 2014. «Performance evaluation of typical approximation algorithms for nonconvex lp-minimization in diffuse optical tomography». *J. Opt. Soc. Am. A*, vol. 31, n° 4, p. 852–862.
- Strangman, G., D. A. Boas, et J. P. Sutton. 2002. «Non-invasive neuroimaging using near-infrared light». *Biological psychiatry*, vol. 52, n° 7, p. 679–693.
- Su, J., Y. Liu, Z. Lin, S. Teng, A. Rhoden, N. Pantong, et H. Liu. 2014. «Reconstructions for Continuous-Wave Diffuse Optical Tomography by a Globally Convergent Method». *Journal of Applied Mathematics and Physics*, vol. 2, p. 204–213.
- Su, J., H. Shan, H. Liu, et M. V. Klivanov. 2006. «Reconstruction method with data from a multiple-site continuous-wave source for three-dimensional optical tomography». *J. Opt. Soc. Am. A*, vol. 23, n° 10, p. 2388–2395.
- Taler, J. 1992. «Nonlinear steady-state inverse heat conduction problem with space-variable boundary conditions». *Journal of heat transfer*, vol. 114, n° 4, p. 1048–1051.
- Tan, H.-P., H.-L. Yi, P.-Y. Wang, L.-M. Ruan, et T. W. Tong. 2004. «Ray tracing method for transient coupled heat transfer in an anisotropic scattering layer». *International journal of heat and mass transfer*, vol. 47, n° 19, p. 4045–4059.
- Tarvainen, T. 2006. «Computational methods for light transport in optical tomography». PhD thesis, University of Kuopio.
- Tarvainen, T., M. Vauhkonen, V. Kolehmainen, et J. Kaipio. 2006. «Finite element model for the coupled radiative transfer equation and diffusion approximation». *International Journal for Numerical Methods in Engineering*, vol. 65, n° 3, p. 383–405.
- Tarvainen, T., M. Vauhkonen, et S. Arridge. 2008a. «Gauss–Newton reconstruction method for optical tomography using the finite element solution of the radiative transfer equation». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 109, n° 17, p. 2767–2778.
- Tarvainen, T., V. Kolehmainen, S. Arridge, et J. Kaipio. 2011. «Image reconstruction in diffuse optical tomography using the coupled radiative transport-diffusion model». *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 112, p. 2600–2608.
- Tarvainen, T., M. Vauhkonen, V. Kolehmainen, S. R. Arridge, et J. P. Kaipio. 2005a. «Coupled radiative transfer equation and diffusion approximation model for photon migration in turbid medium with low-scattering and non-scattering regions». *Physics in medicine and biology*, vol. 50, n° 20, p. 4913.
- Tarvainen, T., M. Vauhkonen, V. Kolehmainen, et J. P. Kaipio. 2005b. «Hybrid radiative-transfer–diffusion model for optical tomography». *Applied optics*, vol. 44, n° 6, p. 876–886.
- Tarvainen, T., M. Vauhkonen, V. Kolehmainen, J. P. Kaipio, et S. R. Arridge. 2008b. «Utilizing the radiative transfer equation in optical tomography». *Piers Online*, vol. 4, n° 6, p. 655–660.
- Tikhonov, A. et V. Arsenin, 1977. *Solution of ill-posed problems*. New-York : John Wiley.



- Tikhonov, A., A. Leonov, et A. Yagola, 1998. *NonLinear Ill-Posed Problems, Vol. 1 and 2*, 1998.
- Torregrossa, M. 2003. «Reconstruction d'images obtenues par tomographie optique dans le proche infrarouge». PhD thesis, Université Louis Pasteur - Strasbourg 1.
- Vogel, C. R. 1996. «Non-convergence of the L-curve regularization parameter selection method». *Inverse problems*, vol. 12, n° 4, p. 535.
- Wang, L. V. et H. Wu, 2012. *Biomedical optics : principles and imaging*.
- Wang, Z., Y. Liu, G. Wang, et L. Sun. 2009. «Elastography method for reconstruction of nonlinear breast tissue properties». *Journal of Biomedical Imaging*, vol. 2009, p. 6.
- Williams, J. R. et K. Amaratunga. 1997. «A discrete wavelet transform without edge effects using wavelet extrapolation». *Journal of Fourier Analysis and Applications*, vol. 3, n° 4, p. 435–449.
- Woodbury, K., 2003. *Inverse Engineering Handbook*. Boca Raton : CRC Press.
- Xu, Y., X. Gu, T. Khan, et H. Jiang. 2002. «Absorption and scattering images of heterogeneous scattering media can be simultaneously reconstructed by use of dc data». *Applied optics*, vol. 41, n° 25, p. 5427–5437.
- Yagola, A., A. Leonov, et V. Titarenko. 2002. «Data errors and an error estimation for ill-posed problems». *Inverse problems in engineering*, vol. 10, n° 2, p. 117–129.
- Ye, J. C., C. A. Bouman, K. J. Webb, et R. P. Millane. 2001. «Nonlinear multigrid algorithms for Bayesian optical diffusion tomography». *Image Processing, IEEE Transactions on*, vol. 10, n° 6, p. 909–922.
- Zacharopoulos, A. D., S. R. Arridge, O. Dorn, V. Kolehmainen, et J. Sikora. 2006. «Three-dimensional reconstruction of shape and piecewise constant region values for optical tomography using spherical harmonic parametrization and a boundary element method». *Inverse Problems*, vol. 22, n° 5, p. 1509.
- Zhang, X. 2014. «Instrumentation in diffuse optical imaging». In *Photonics*. p. 9–32. Multidisciplinary Digital Publishing Institute.
- Zhang, X., K. Liu, et J. Liu. 2006. «The wavelet multiscale method for inversion of porosity in the fluid-saturated porous media». *Applied mathematics and computation*, vol. 180, n° 2, p. 419–427.
- Zhao, J., T. Liu, et S. Liu. 2014. «Identification of space-dependent permeability in nonlinear diffusion equation from interior measurements using wavelet multiscale method». *Inverse Problems in Science and Engineering*, vol. 22, n° 4, p. 507–529.
- Zhu, W., Y. Wang, Y. Deng, Y. Yao, et R. L. Barbour. 1997a. «A wavelet-based multiresolution regularized least squares reconstruction approach for optical tomography». *Medical Imaging, IEEE Transactions on*, vol. 16, n° 2, p. 210–217.
- Zhu, W., Y. Wang, Y. Yao, J. Chang, H. L. Graber, et R. L. Barbour. 1997b. «Iterative total least-squares image reconstruction algorithm for optical tomography by the conjugate gradient method». *J. Opt. Soc. Am. A*, vol. 14, n° 4, p. 799–807.
- Zhu, W., Y. Wang, et J. Zhang. 1998. «Total least-squares reconstruction with wavelets for optical tomography». *J. Opt. Soc. Am. A*, vol. 15, n° 10, p. 2639–2650.